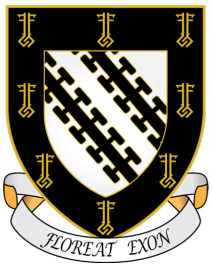
The Dynamical Approach to Relativity as a Form of Regularity Relationalism

---

Syman Stevens

Exeter College, Oxford University



Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Trinity Term, 2014

The Dynamical Approach to Relativity as a Form of Regularity Relationalism

Syman Stevens

Exeter College, Oxford · Trinity Term, 2014

[Abstract]

This thesis investigates the interplay between explanatory issues in special relativity and the theory's metaphysical foundations. Special attention is given to the 'dynamical approach' to relativity, promoted primarily by Harvey Brown and collaborators, according to which the symmetries of dynamical laws are explanatory of relativistic effects, inertial motion, and even the Minkowskian geometrical structure of a specially relativistic world.

The thesis begins with a review of Einstein's 1905 introduction to special relativity, after which brief historical introductions are given for the standard 'geometrical' approach to relativity and the unorthodox 'dynamical' approach. After a critical review of recent literature on the topic, the dynamical approach is shown to be in need of a metaphysical package that would undergird the explanatory claims mentioned above. It is argued that the dynamical approach is best understood as a form of relationalism—in particular, as a relativistic form of 'regularity relationalism', promoted recently by Nick Huggett. According to this view, some portion of a world's geometrical structure actually supervenes upon the symmetries of the best-system dynamical laws for a material ontology endowed with a primitive *sub*-metrical structure.

To explore the plausibility of this construal of the dynamical approach, a case study is carried out on solutions to the Klein–Gordon equation. Examples are found for which the field values, when purged of all spatiotemporal structure but their induced topology, are still arguably best-systematized by the Klein–Gordon equation itself. This bolsters the plausibility of the claim that some system of field values, endowed with mere sub-metrical structure, might have as its best-systems dynamical laws a (set of) Lorentz-covariant equation(s), on which Minkowski geometrical structure would supervene.

The upshot is that the dynamical approach to special relativity can be defended as what might be called an ontologically and ideologically relationalist approach to Minkowski spacetime structure. The chapters refer regularly to three appendices, which include a brief introduction to topological and differentiable spaces.

Document Statistics

Text in Body and Footnotes: 73,149

Text in Captions: 681

Text in References: 6,454

In-line Math Terms: 1,237

Lines of Displayed Math: 31

Number of Figures: 37

Total Word Count: 75,067 *

*This total includes captions, footnotes, and in-line citations. It counts each in-line mathematical term as one word. It does not include displayed mathematics or the References section.

Nomenclature

LST	Leibnizian Spacetime
GST	Galilean Spacetime
MST	Minkowski Spacetime
(...)	Clarifications in body and footnotes; Internal cross references
[...]	Citations; Clarifications in quotations
p. , pp.	Page, Pages
L , R	Left-, Right-hand column of page
fn. , Eqn. , Fig.	Footnote, Equation, Figure
cf. , ff	Compare with, And following pages
i.e. , e.g.	That is, For example
Ch. , § , Appx.	Chapter, Section, Appendix
P	Set of (field-)points
p	Member of set P
$\mathcal{T}$	Class of sets, encoding topological structure
T	Set member of $\mathcal{T}$, i.e. an 'open set' in $\mathcal{T}$
$\langle P, \mathcal{T} \rangle$	Space of elements of P with topological structure $\mathcal{T}$
$\langle P, \eta \rangle$	Space of elements of P with Minkowski 'metrical' structure η
$\mathbb{R}^4$	Set of ordered quadruples of $\mathbb{R}$, with standard Euclidean metric
$\mathcal{R}^4$	Standard Euclidean topology of (induced by) $\mathbb{R}^4$
R	Open set in $\mathcal{R}^4$
$\underline{\circledcirc}$	Open subset
$\underline{\bullet}$	Closed subset
$a\mathrm{R}b$	That a stands in relation R to b
$[a]_{\mathrm{R}}$	Equivalence class of a with respect to (equivalence) relation R
P/R	Quotient of set P with respect to (equivalence) relation R
$\bar{v}$	Three-vector
$\vec{v}$	Four-vector
$f(x)$	That f is a function of x
$f[a]$	Evaluation of f at $x = a$
$\phi : P \to Q$	That ϕ is a mapping from P into Q
$\phi[T]$	Evaluation, or image, of ϕ acting on set T

List of Figures

Table of Contents

Introduction:
Explanation and Geometry in Special Relativity

The subject of this thesis falls under the intersection of explanation and interpretation in spacetime theories. In particular, this thesis investigates the interplay between explanatory issues in special relativity and the theory's metaphysical foundations. That is, in an attempt to answer questions such as "What explains relativistic phenomena?", and especially "What is the arrow of explanation between Lorentz covariance and Minkowski spacetime structure?", this thesis gives careful attention to the underlying metaphysical details of rival interpretations of special relativity. Special attention is given to one particular unorthodox interpretation, according to which the symmetries of dynamical laws are explanatory of relativistic effects, inertial motion, and even the Minkowskian geometrical structure of a specially relativistic world. This interpretation is better known as the 'dynamical approach' to special relativity, and has been promoted primarily by Harvey Brown and his collaborators, most notably Oliver Pooley.[1]

While there are some metaphysical claims implicit in Brown's work, they are not central to his agenda. For whereas Brown defends the explanatory priority of dynamical symmetries over spacetime geometry, he does not explicitly defend any particular conception of natural laws, or any particular position in the substantivalist–relationalist spectrum of approaches to spacetime. This thesis does just that. In the chapters outlined below, the dynamical approach to special relativity is introduced, defended from a number of recent criticisms, and furnished with a metaphysical package that undergirds the explanatory claims mentioned above. The end result is what might be called an ontologically and ideologically relationalist approach to the geometrical structure of Minkowski spacetime.[2] And while there may be other ways to understand and defend the

[1] Brown's independent work and collaborations with others are introduced in Chapter 1.

[2] Throughout, the terms 'ontology' and 'ideology' are used as by Quine [1951, p. 11, p. 14]. The ontology of a theory comprises the objects over which its bound variables range (i.e. the elements of the domain of its models), and the ideology of a theory amounts to "what ideas can be expressed in it" (i.e. what properties are taken as primitive and used in predication). What is meant here by 'ontological and ideological relationalism' is fleshed out in the final paragraphs of Chapter 2.

dynamical approach to special relativity, this construal is shown to make the dynamical approach defensible despite recent criticisms, and also to make the explanatory claims of the dynamical approach more plausible than they might otherwise seem.

Chapter 1 of this thesis is introductory. It begins with a brief review of Einstein's 1905 introduction to special relativity, giving special attention to the explanatory issues that are the focus of the following chapters. A brief historical introduction is also given for the 'geometrical' approach to relativity, against which the 'dynamical' approach is finally introduced.

Chapter 2 goes on to engage directly with the recent literature, and focuses primarily upon work introduced in Chapter 1. First and foremost, it is shown that recent work on the arrow of explanation between Lorentz covariance and Minkowski spacetime structure has suffered from a difference in various authors' interpretations of Einstein's comments regarding the (lack of) explanatory power of his 1905 derivation of the Lorentz transformations. Then, putting Einstein's framework behind us, other recent criticisms of the dynamical approach are introduced. It is the collective aim of the following few chapters to give an adequate response to them.

Chapter 3 takes a first step in introducing one compelling construal of the dynamical approach, particularly regarding the claim that the symmetries of dynamical laws can account for Minkowski spacetime structure. It is shown that this claim gives the dynamical approach a unique position in the well-rehearsed empiricist–conventionalist debate on spacetime geometry. Indeed, the dynamical approach finds its most natural home somewhere in the vicinity of Einstein's 'practical' approach to geometrical structure, despite the fact that Einstein seems to present those ideas as a form of empiricism, which the dynamical approach is not. Some common features of the dynamical approach and Einstein's 'practical' approach to geometry are then traced through some early work on special relativity, showing that this construal of the dynamical approach is not an entirely new perspective on relativity and geometry. Finally, this construal of the dynamical approach is shown to be consistent with Brown's work on "the physics of coordinate

transformations" in his book, *Physical Relativity*.

Chapter 4 seeks out a metaphysical package to undergird the explanatory claims of the dynamical approach, as it was described in the preceding chapter. It is argued that the dynamical approach is best understood as a particular form of relationalism—namely, a relativistic form of so-called 'regularity relationalism'. According to this view, some portion of a specially relativistic world's geometrical structure actually supervenes upon the symmetries of the best-system dynamical laws for a material ontology endowed with a primitive *sub*-metrical structure. A particularly simple choice of sub-metrical structure is introduced—namely, the standard Euclidean topology—so that the stage is set for a search of supportive examples in which Minkowski spacetime structure could so supervene. That search is carried out in the following chapter.

Chapter 5 carries out a case study on regularity relationalism for Minkowski spacetime structure. It introduces planewave solutions to the Klein–Gordon equation and shows how, when purged of all but their topological structure, the values can still be systematized by the same equation. And while it may be that the simplest such solutions can also be systematized by equally simple dynamical laws of distinct covariance groups, more complex systems are described which, while still systematized by the Klein–Gordon equation, are not so easily systematized by other equations. Given that the Klein–Gordon equation is Lorentz covariant, so that upon its symmetries would supervene Minkowskian geometrical structure, this shows that the explanatory claims of the dynamical approach, when the latter is construed as a form of regularity relationalism, are much more plausible than they might at first appear.

Chapter 6 gives some concluding remarks. First, some potential concerns or criticisms of this thesis's construal of the dynamical approach are addressed, as are some potential roadblocks in taking the dynamical approach beyond the context of a simple scalar field in special relativity. After this, another overview of the chapters and their main points is given, this time supplemented with suggestions for further research. Here follow the appendices, which cover some technical details and tangential issues.

Appendix A provides an introduction to the amorphous geometrical structures of topological spaces and differentiable manifolds. An emphasis is placed on those concepts most relevant to the preceding chapters. For those unfamiliar with the mathematical formalism of point-set topology, it would be good to have read this appendix before reading Chapters 4 and 5.

Appendix B is a supplement to the material in Chapter 5. It is shown that applying these techniques to systems of more than one physical field does not require the introduction of a primitive coincidence relation between field-points that constitute a 'material event'. Rather, by taking a particular form of topological structure as primitive, coincidence may be reduced to topological indiscernibility.

Appendix C carries out the same project as in Chapter 5, but takes a much richer sub-metrical structure as primitive—namely, Euclidean affine structure. In that context, certain simplifications are possible that were not available when Euclidean topological structure was taken as primitive. With those simplifications at one's disposal, accounting for Minkowski spacetime structure along the lines described above becomes even more straightforward.

1 Rival Approaches to Relativity:
The Geometrical and Dynamical Approaches to Special Relativity

This thesis focuses primarily upon a somewhat unorthodox interpretation of special relativity. In what follows, that approach and its closest rival are referred to as the 'dynamical' and 'geometrical' approaches to special relativity, respectively.[1] In this section, each is introduced by way of a brief literature review. Of course, an introduction to these two rivals would benefit from a review of the work that introduced special relativity: Einstein's 1905 article, *On the Electrodynamics of Moving Bodies*. That article is introduced in §1.1. The geometrical and dynamical approaches to special relativity are then introduced in §1.2 and §1.3, respectively.

1.1 Einstein, 1905: *On the Electrodynamics of Moving Bodies*

Einstein's original presentation of special relativity [1905*b*] brought about what was arguably the most significant development in spacetime physics since Newton's *Principia Mathematica*, published more than two centuries earlier [1687]. There is enough to say about the development of the ideas in Einstein's paper—and also about their immediate and long-term impact—to fill more pages than allotted here. Indeed, plenty has already been said elsewhere.[2] Here, the focus is limited to the uniqueness of Einstein's approach to giving a physical description of electrodynamical phenomena.

With that aim, this section begins with a review of the context in which *On the Electrodynamics of Moving Bodies* was published. Then, after outlining the structure of the article itself, it ends with an overview of Einstein's thoughts on precisely what 'kind' of theory he was promoting in it. It is shown in later sections that those thoughts

[1] Some authors refer to these as the dynamical and geometrical 'interpretations'. And some, like Nerlich [1994*b*] and Craig [2001; 2008], refer to the latter as the 'spacetime' interpretation.

[2] For instance: Stachel and Janssen [2004] outline work on electromagnetic phenomena in ether-theories in the century leading up to 1905; Stachel [2002, §IV] unravels the development of Einstein's thinking in the years directly preceding 1905; and Miller [1981] complements both of these works by giving an authoritative analysis of the emergence of special relativity in the years surrounding 1905. These and other sources, including work by Zahar [1989, Chs. 1–3] and Craig [2001, Chs. 1–2], are referenced in the body of this section.

have had a significant impact on others' understanding of how such electrodynamical phenomena ought to be explained.

Ether and the Contraction Hypothesis

At the start of the twentieth century, decades of progress in both theory and experiment stood in support of a wave-theory of light.[3] Although the particle-theory of light presented in Isaac Newton's *Opticks* [1704] had dominated mainstream thinking for the century that followed, 19th-century work like Thomas Young's study of interference patterns [1804], and Augustin Fresnel's calculations on diffraction and polarization [1870], had brought the wave-theory back into position as the dominant view.

At the time, it was thought that, like waves in more familiar contexts, light passed through some medium—the 'luminiferous ether'—at a constant speed that was independent of the speed of the light's source. The details of that propagation remained controversial until the publication of James Clerk Maxwell's groundbreaking work, *Treatise on Electricity and Magnetism* [1873].[4] Maxwell's theory introduced electromagnetic waves, which propagated at a constant rate that was later found to be the speed of light.

Maxwell's theory was introduced at a time when Newtonian kinematics was still in vogue.[5] That is to say, electrodynamics was first received by a community of physicists who shared the assumption that the Cartesian coordinates (x', y', z', t') for a frame of reference moving with respect to some 'rest' frame with coordinates (x, y, z, t) would be determined by a Galilean boost.

$$\begin{aligned} x' &= x - vt & \qquad y' &= y \\ t' &= t & \qquad z' &= z \end{aligned} \tag{1}$$

It was also believed that the relative velocity (v_{rel}) of an object with respect to some observer would be determined by the Galilean velocity addition rule, according to which

[3] For a historical overview of the wave- and particle-theory of light, see recent work by Stark [2013].

[4] Wave-propagation theories actually date back to before Newton's *Opticks*. For example, Descartes, Hooke, and Huygens all proposed descriptions of light as a wave [Stark 2013].

[5] Throughout this thesis, the term 'kinematics' is used to refer to the study of motion without attention to forces or interactions between material bodies—i.e. to the study of force-free equations.

the velocity of the observer (v_{obs}) and that of the object (v_{obj}) were both measured with respect to one and the same frame of reference.

$$v_{rel} + v_{obs} = v_{obj}$$

So, given that the speed of electrodynamic waves with respect to the ether (v_{obj}) was constant in Maxwell's dynamics, it should have been possible to determine the relative velocity of some reference frame with respect to the ether (v_{obs}) by measuring the speed of light with respect to that frame (v_{rel}). Indeed, attempts at measuring some effect of the Earth's motion with respect to the ether date back at least as far as James Bradley's test on starlight aberration [1727].[6] And with Maxwell's dynamics in hand, and the increasing accuracy of experimental devices at the time, the stage was set for a successful measurement to be made at last. Albert Michelson was soon commissioned to perform an experiment along these lines [1881], which failed to produce a positive result. Experimental errors were later found to be the cause of failure, and Michelson performed another, more accurate experiment with Edward Morley [1887]. But it also, to much greater fame, failed to produce a positive result for the Earth's motion relative to the ether.

Following these events, two physicists independently proposed a contraction hypothesis as an explanation of Michelson and Morley's null result. One was George Francis FitzGerald, who, in a letter to *Science* magazine, suggested that the null result could be explained by assuming that the arms of the apparatus had actually contracted "by an amount depending upon the square of the ratio of their velocity to that of light" [1889].[7] Recalling earlier work by Oliver Heaviside [1888], which entailed distortions in

[6] Early attempts at finding observable effects of the Earth's motion with respect to the ether also include Arago's 1810 test for a violation of Snell's law, and Hippolyte Fizeau's 1851 experiment on interference patterns in moving water. These and others are discussed by Stachel and Janssen [2004], who also explain how the null results of these experiments were later accounted for by theories of a partially-dragged ether. The experiments of Michelson and Morley were unique in being sensitive to second-order effects, so that the same explanations could not account for their null results. This is also discussed further by Brown [2005*b*, pp. 2–3, pp. 42–5].

[7] This and other of FitzGerald's correspondences are discussed further by Brown [2005*b*, Ch. 1]. Brown also notes that FitzGerald (and Lorentz, introduced below) actually proposed more general 'deformation' hypotheses, which may have even involved length expansion [pp. 3, 53].

the electrical field of a charged particle moving with respect to the ether, FitzGerald wondered if the same might be true for all inter-molecular forces. His letter continued as follows.

> We know that electric forces are affected by the motion of the electrified bodies relative to the ether, and it seems a not improbable supposition that the molecular forces are affected by the motion, and that the size of a body alters consequently. [FitzGerald 1889, p. 390]

FitzGerald offered no suggestions as to why or how the apparatus's motion through the ether would effect such a contraction, but of course he was right—such a contraction would indeed account for the null result.

The other was Hendrik Lorentz. Between the time of Michelson's first experiment and his later collaboration with Morley, Lorentz had discovered that the description of moving charges in Maxwell's theory would preserve its form when expressed in terms of new coordinates [1892*a*]. The coordinate transformations were very near to what are now known as the Lorentz-transformations, and which are written in standard configuration as follows.

$$x' = \frac{x - vt}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \qquad y' = y \tag{2}$$

$$t' = \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \qquad z' = z$$

At the time, Lorentz made use of the coordinate transformations as mere mathematical devices that simplified the description of moving charges. Working within the framework of Newtonian kinematics, he did not take them to be the true coordinate transformations between inertial frames. However, he took an important step in that direction soon thereafter. Following a similar train of thought as FitzGerald, Lorentz promoted what he called the 'Molecular Force Hypothesis'—namely, the hypothesis that the mathematical description of all molecular forces might share this same feature.

> Now, some such change in the length of the arms of Michelson's first experiment, and in the dimensions of the slab in the second one is so far as I

> can see, not inconceivable. What determines the size and shape of a solid body? Evidently the intensity of the molecular forces; any cause which would alter the latter would also influence the shape and dimensions. Nowadays we may safely assume that electric and magnetic forces act by means of the intervention of ether. It is not far-fetched to suppose the same to be true of the molecular forces. [Lorentz 1892*b*, p. 221]

In the same work, Lorentz went on to show that if the Molecular Force Hypothesis were true, the resulting contraction in material bodies moving through the ether would be precisely the amount needed to account for the null result of Michelson and Morley's experiment.[8]

In sum, FitzGerald and Lorentz proposed similar contraction hypotheses in order to account for Michelson and Morley's null result. This hypothesis, often referred to as the 'generalized contraction hypothesis' amounts to the assumption that all forces—not just electromagnetic forces—governing material bodies are Lorentz-covariant.[9] And because this concept plays such a critical role in the following sections, it deserves to be spelled out further. Dynamical laws are deemed 'Lorentz-covariant' when they preserve their form under any combination of the coordinate transformations outlined in Equation 2, along with the four one-dimensional 'translation' transformations and the three 'spatial rotation' transformations. Together, these form a group—the 'Poincaré group'—which serves as the 'covariance group' of a Lorentz-covariant equation.[10] And so in what follows, a Lorentz-covariant equation is one which preserves its form under coordinate transformations of the Poincaré group.[11]

---

[8] The molecular force hypothesis was also discussed at greater length in Lorentz's next major work [1895]. That work is discussed at length by Zahar [1989, §§2.2–2.5, esp. pp. 64ff]. It should be noted that Lorentz originally suggested a more general deformation that involved length contraction as only one of its possible effects. This is also discussed by Brown and Pooley [2001, p. 258] and by Brown [2005*b*, pp. 52–3].

[9] Or nearly so—as mentioned above, Lorentz's original coordinate transformations were not exactly those shown in Eqn. 2. Rather, Lorentz included an extra factor of $\gamma = 1/\sqrt{1-(v/c)^2}$ in the transformation of the time coordinate. And while those transformations would account for length contraction, they would not account for time dilation; Lorentz did not recognize the full covariance of Maxwell's equations until after Einstein derived the Lorentz transformations in 1905 [Zahar 1989, §§2.2–2.3; Craig 2001, pp. 17–8; Brown 2005*b*, §1.4].

[10] The Poincaré group is also known as the 'inhomogeneous Lorentz group'. Unfortunately, it is sometimes referred to as the 'Lorentz group', as in Friedman's otherwise helpful discussion of it [1983, p. 136].

[11] This has been said with respect to an equation's 'standard formulation'. Although any theory can be expressed in generally covariant form, its standard formulation preserves its form under the transformations

Of course, FitzGerald's suggestion did not even approach this kind of mathematical detail. His idea was motivated by the aforementioned work of Heaviside, but lacked any further theoretical justification. As a result, it has since been described, rather unfairly, as "the brilliant baseless guess of an Irish genius".[12] Lorentz, however, did develop independent theoretical justification for the hypothesis by working toward empirically adequate dynamical laws for the interactions between material bodies and the point-like parts of the luminiferous ether.[13] This project would continue to occupy Lorentz until after Einstein's introduction to special relativity which, it is shown below, came out of Einstein's own skepticism that anyone could succeed in the kind of project that Lorentz had set before him.[14]

Einstein's Introduction to Special Relativity

At the turn of the century, Lorentz's program was arguably the most promising attempt to reconcile Maxwell's dynamics with failed attempts to measure the Earth's speed with respect to the ether. However, a novel approach to these matters quickly took center stage with the publication of Einstein's *On the Electrodynamics of Moving Bodies* [1905*b*]. To be clear: this is not to say that Einstein was trying his own hand at explaining the null result of the Michelson–Morley experiment. For although Einstein claims to have read, before 1905, work of Lorentz's in which the experiment is discussed [1895], it is not clear that the experiment played any significant role in Einstein's thinking at the time.[15]

---

of its own covariance group. The covariance group of a theory's standard formulation is often referred to as the 'symmetry group' or 'invariance group' of the theory itself, as it is by Friedman [1983, pp. 56–61]. Brown makes some related comments [2005*b*, p. 75, fn. 20].

[12] Brown [2005*b*, p. 49], quoting Glazebrook [1928]. Brown offers further discussion [2005*b*, p. 49, fn. 26].

[13] Lorentz's theory of corresponding states involved an immobile ether, in which electric and magnetic fields could be formed by any charged material particles located in it. The early success of this theory was in explaining the results of the experiments mentioned in footnote 6 on page 7. Janssen [2002*b*] and Stachel and Janssen [2004] give a comprehensive overview, while Zahar [1989] gives a more technical analysis.

[14] Despite his efforts to defend himself, Lorentz's approach gained a reputation for being *ad hoc*, just as FitzGerald's hypothesis did [Brown 2005*b*, p. 41, pp. 53–4; Brown 2005*a*, p. 43]. Even by 1908, Hermann Minkowski would describe the hypothesis as "extremely fantastical, for the contraction is not to be looked upon as a consequence of resistances in the ether, or anything of that kind, but simply as a gift from above—as an accompanying circumstance of the circumstance of motion" [1909, p. 81]. Arguments that Lorentz's theory was not *ad hoc* are given by Craig [2001, pp. 13–8], Zahar [1989, §2.1], and Janssen [2002*b*, pp. 431ff].

[15] Brief discussions, with references, are provided by Janssen [2002*b*, pp. 431–2] and Craig [2001, pp. 21–3].

But regardless of Einstein's motivations, his work provided a new account of relativistic phenomena like length contraction. For whereas Lorentz, working within the framework of Newtonian kinematics, proposed a physical contraction resulting from material bodies' interaction with the ether, Einstein brought about a change in kinematics by treating both physical descriptions of material bodies—those in the original and Lorentz-transformed coordinates—as equally valid. Unlike Lorentz, then, Einstein denied there being any frame of reference with respect to which the physical description was any truer than it was in others, regardless of whether that be a frame of absolute rest or simply the rest frame of the ether.

This train of thought is clear in the way that Einstein opens the paper.[16]

> It is known that Maxwell's electrodynamics—as usually understood at the present time—when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena. Take, for example, the reciprocal electrodynamic action of a magnet and a conductor. The observable phenomenon here depends only on the relative motion of the conductor and the magnet, whereas the customary view draws a sharp distinction between the two cases in which either the one or the other of these bodies is in motion. [...]
>
> Examples of this sort, together with the unsuccessful attempts to discover any motion of the earth relatively to the "light medium," suggest that the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest. [Einstein 1905*b*, p. 37]

In order to do away with these asymmetries, Einstein does away with the concept of absolute rest. And to motivate such a change in kinematics, Einstein gives an insightful analysis of the nature of time via a discussion on how spatially separated clocks can be synchronized. But as his ideas were presented in 1905, these insights into the nature of time were not presented at the outset, but rather as a first step in accounting for two observed regularities that he first raises to the status of postulates.

1. The relativity principle: "that the same laws of electrodynamics and optics will be valid for all frames of reference for which the equations of mechanics hold good."

[16] Einstein would later explain that the idea of treating all frames equally came to him by reflecting upon the philosophical works of Hume and Mach [Einstein 1949*a*, p. 53].

2. The light postulate: "that light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body."
[Einstein 1905*b*, pp. 37–8]

Einstein's genius is revealed in showing that these two seemingly incompatible postulates can be reconciled once careful attention is given to our concept of time. He begins by defining the 'time' of any event to be "the reading simultaneous with the event of a clock at rest and located at the position of the event, this clock being synchronous, and indeed synchronous for all time determinations, with a specified clock at rest" [p. 40], where 'synchronous' refers to the following definition of simultaneity by the transmission of light beams.

> [We] establish *by definition* that the "time" required by light to travel from A to B equals the "time" it requires to travel from B to A. Let a ray of light start at the "A time" t_A from A towards B, let it at the "B time" t_B be reflected at B in the direction of A, and arrive again at A at the "A time" t'_A. In accordance with definition the two clocks synchronize if $t_B - t_A = t'_A - t_B$. [Einstein 1905*b*, p. 40]

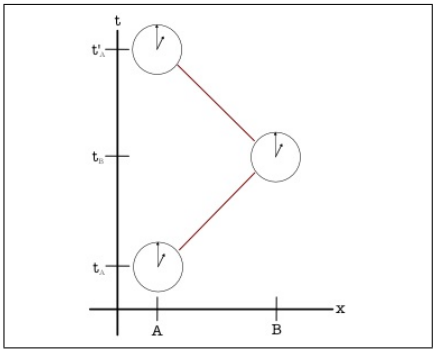


Figure 1: Einstein's clock synchrony.

And with time so defined, Einstein shows that his two postulates entail a difference in measurements of the length of a rod, or of the elapsed time on a clock, for two observers in relative motion [§2]. He goes on to show that the proper coordinate transformations between two inertial frames are in fact the Lorentz transformations [§3], and then gives a physical interpretation of this fact as the now-familiar phenomena of length contraction and time dilation [§4].[17]

Now, in deriving the Lorentz transformations from his two postulates, Einstein effectively showed that Lorentz's contraction hypothesis and the corresponding dilation

[17] For commentary on Einstein's steps in deriving the Lorentz transformations from his two postulates (and other assumptions, like the isotropy of space), see any of the sources in footnote 2 on page 5. Brown also offers some discussion [2005*b*, §§5.3–5.4].

in clocks come out of his understanding of time. He would later explain how he came to recognize that this was the case.[18]

> It became evident, however, astonishingly, that it was only necessary to comprehend the concept of time in a sufficiently exact way in order to get over the difficulty just stated. [That difficulty being the need for a contraction hypothesis.] It required only the realization that one can define simply as "time" an auxiliary quantity introduced by H. A. Lorentz which he called "local time." If one sticks to the definition of time just indicated, then the basic equations of the Lorentz theory do conform to the principle of relativity, if one only replaces the transformation equations given above (the Galilean transformations) by ones that correspond to the new concept of time. The hypothesis of H. A. Lorentz and Fitzgerald then comes out as a compelling result of the theory. [Einstein 1907*b*, as quoted by Klein [1982]]

And so unlike Lorentz, Einstein did take the Lorentz transformations to be the true coordinate transformations between inertial frames. But despite Einstein's emphasis in the preceding quotation on the role played by his conception of time, his work in 1905 was based primarily upon the two postulates introduced above. His definition of time appears only as a sort of precondition of the two postulates, which are the real starting point of the theory.

Einstein's Principle-Theory Compromise

From the brief outline above, it should be clear that whereas Lorentz's work focused upon the details of the dynamical laws governing material objects, Einstein focused instead on the kinematical and dynamical implications of reconciling his two postulates. But as is clear in his Autobiographical Notes, Einstein later worried that in doing so, he had effectively taken measuring rods and time-keeping devices as privileged entities.[19]

> One is struck that the theory (except for the four-dimensional space) introduces two kinds of physical things, i.e., (1) measuring rods and clocks, (2) all other things, e.g., the electromagnetic field, the material point, etc. This, in

[18] In his celebrated lecture, *Space and Time*, Minkowski makes a similar point [1909, p. 82]. Poincaré's very similar and slightly earlier ideas are discussed by Brown [2005*b*, §4.6].

[19] This is discussed at greater length by Brown [2005*b*, §1.3, §5.5].

> a certain sense, is inconsistent; strictly speaking measuring rods and clocks would have to be represented as solutions of the basic equations (objects consisting of moving atomic configurations), not, as it were, as theoretically self-sufficient entities. [Einstein 1949*a*, p. 59]

In other passages of the notes, Einstein is clear about his reasons for treating rods and clocks the way he did. Given then-recent developments in quantum theory, including his own work on electromagnetic radiation [1905*a*], Einstein doubted the possibility of succeeding in a project like Lorentz's. Instead, Einstein recognized the value in using principles as heuristics in theory development, and saw thermodynamics as a positive example.[20]

> Reflections of this type made it clear to me as long ago as shortly after 1900, i.e., shortly after Planck's trailblazing work, that neither mechanics nor electrodynamics could (except in limiting cases) claim exact validity. By and by I despaired of the possibility of discovering the true laws by means of constructive efforts based on known facts. The longer and the more despairingly I tried, the more I came to the conviction that only the discovery of a universal formal principle could lead us to assured results. The example I saw before me was thermodynamics. The general principle was there given in the theorem: the laws of nature are such that it is impossible to construct a *perpetuum mobile* (of the first and second kind).
>
> [Einstein 1949*a*, pp. 51–3; *cf.* pp. 45–7]

The result is what Einstein called a 'principle-theory'.[21] As in Einstein's 1905 work, a principle-theory is one for which the starting point is some set of well-established empirical regularities, and the end result of which is some kind of restriction on the range of dynamical laws with respect to which material objects' behavior would be consistent with those regularities. When contrasted with so-called 'constructive' theories à la Lorentz, the starting point of which is the dynamics that governs whatever primitive

[20] This fact is also outlined by Brown and collaborators [Brown and Pooley 2001, pp. 4–5; Brown 2005*b*, §5.1; Brown and Timpson 2006, §2]. Similar sentiments are expressed by Einstein elsewhere [1919].

[21] In addition to the several sources listed by Brown and Pooley [2006, §3], Arthur Miller also cites a relevant letter from Einstein to Maurice Solovine, in which Einstein calls his 1905 work a principle-theory [1981, pp. 133–4]. For more on Einstein's dissatisfaction with taking this approach, see the discussion and references provided by Brown [2005*b*, §1.6, §5.2] or by Brown and Pooley [2001]. Another relevant letter to Sommerfeld is discussed by Brown and Timpson [2006, pp. 31–2].

entities constitute the system under observation, the principle–constructive distinction raises a number of questions about the structure of physical theories and the quality of explanations that they provide.[22] That distinction and those issues are revisited below (Ch. 2). For now, suffice it to say that whatever shortcomings they may have, principle-theories are clearly of use as heuristic aides in theory development, which was in fact Einstein's goal in 1905.[23] Indeed, Einstein later claimed that the take-home lesson from his 1905 paper was effectively a restriction on the form of the laws of any constructive theory.[24]

> The universal principle of the special theory of relativity is contained in the postulate: The laws of physics are invariant with respect to Lorentz-transformations (for the transition from one inertial system to any other arbitrarily chosen system of inertia). This is a restricting principle for natural laws, comparable to the restricting principle of the non-existence of the *perpetuum mobile* which underlies thermodynamics.
>
> [Einstein 1949*a*, pp. 57–9]

In this way, Einstein's main result was in some ways quite similar to a generalized contraction hypothesis. But whereas FitzGerald's and Lorentz's hypotheses were at worst *ad hoc* and at best motivated by an incomplete constructive theory of electrons' interactions with the ether, Einstein's principle-theory approach gave his restricting principle as firm a foundation as the laws of thermodynamics.[25]

1.2 The Geometrical Approach to Special Relativity

We have seen already that there was nothing particularly 'geometrical' about Einstein's work in 1905. At its heart were two postulates whose reconciliation called for a change of kinematics, which entailed a restriction on the structural form of dynamical laws. Nonetheless, geometry came to play a major role in future work on relativity, and

[22] Don Howard calls this Einstein's most original contribution to the philosophy of science [2010, p. 35].
[23] Einstein describes his work as a "heuristic principle" in reply to a note by Ehrenfest [Einstein 1907*a*]. Einstein's reply is also discussed by McCormmach [1970, pp. 488–9] and by Klein [1967, p. 515].
[24] This point is discussed further by Brown [2005*b*, §5.1].
[25] Martin Klein makes a similar point [1982, p. 366]. See also footnote 14 on page 10, above.

especially in the way that special relativity has come to be taught. The development of that role is outlined below, and sets the stage for introducing what has come to be known as the 'geometrical' approach to relativity. To be clear, the aim of this section is to introduce, but not necessarily to critique, the geometrical approach. A more critical review of some of the works introduced here is given in Chapter 2.

Before beginning, the following clarification should be made: the preceding and following references to (Minkowski) spacetime structure, or (Minkowskian) geometrical structure, entail no implicit adherence to the existence of 'substantival' spacetime as an independently existing entity. Unless otherwise specified, every mention of spacetime structure should be taken to refer to the systematic forms or patterns in the spatiotemporal arrangement of the world's fundamental ontology, regardless of whether that ontology includes spacetime and its parts. Likewise, every mention of spacetime geometry should be taken to refer to those forms or patterns in the specific case that they conform to the properties of some geometrical structure.[26]

Minkowski's Geometrization of Special Relativity

It was mentioned above that Einstein brought about a kinematical change to spacetime physics. Another way to understand this change is to say that Einstein's work brought about a change in the spacetime geometry employed by physics. That is, the assumption of Newtonian kinematics is that objects behave in accord with the symmetries of the corresponding geometrical structure of Newtonian spacetime—or more accurately, of Galilean spacetime.[27] And when Einstein showed that the covariance group of the world's natural laws is not the Galilean group, but rather the Poincaré group, the revised kinematics amounts to the assumption that objects behave in accord with the symmetries

[26] It becomes clear below that while some advocates of the 'geometrical approach' defend the existence of substantival spacetime, others do not, and not all of the former do so for reasons related to the explanatory issues at hand. Michel Janssen, whose work is introduced below [2009], defends a geometrical approach to special relativity in both the context of substantivalism and that of relationalism—the latter being the view according to which substantival spacetime is not included in the fundamental ontology.

[27] The differences between Neo-Newtonian and Galilean spacetime are rehearsed by Earman [1989, Ch. 2]. Here, I ignore the additional symmetries of Newtonian gravity, which make the theory fit best in Newton–Cartan formulation [Pooley 2013, §6.1.1; 2014, Ch. 4]. See also footnote 3 on page 94.

of a different geometrical structure.[28]

In 1909, Hermann Minkowski provided an exposition of this other geometrical structure [1909]. Essentially, Minkowski recognized that the Lorentz transformations could be visualized as a particular kind of rotation in a non-Euclidean space now known as Minkowski spacetime.[29] For objects to behave in accord with the symmetries of this geometrical structure would mean that space and time no longer enjoy the sort of independence that they did in classical physics, but of course such ideas were already on the table given the experimental verification of the Lorentz transformations. In fact, it is in those experimental results that Minkowski finds justification for his work.

> The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality. [Minkowski 1909, p. 75]

But rather than starting with the experimental physics he mentions, Minkowski decides instead to give purely mathematical reasons for accepting that the world may have some non-Euclidean spatiotemporal structure.

Minkowski begins this process by introducing the coordinates x, y, z, and t of what was for the first time being called a four-dimensional 'world'. After then introducing two components of the Galilean group, which are now referred to as static shifts and kinematical shifts, Minkowski points out that the latter allow alternative permissible coordinatizations that do not preserve the orthogonality between the time axis and spatial hypersurfaces.

[28] Here and elsewhere, any reference to the symmetries of a geometrical structure should be taken to mean those transformations that preserve the forms and patterns mentioned in the introduction to this section. Alternatively, these symmetries can be understood along the lines of Earman's 'space-time symmetries', which are defined in footnote 40 on page 24.

[29] For a more mathematically formal approach to Minkowski spacetime structure, see the introductions by Naber [2012], or by Giulini [2006, §5; 2010]. A brief overview of the Lorentz transformations as hyperbolic rotations is given by Friedman [1983, pp. 135ff], Catoni *et al.* [2008, §4.1.2], and Catoni *et al.* [2011, Ch. 2].

> But the second group [kinematical shifts] means that we may—also without changing the expression of the laws of mechanics—replace x, y, z, t by $x - \alpha t$, $y - \beta t$, $z - \gamma t$, t with any constant values of α, β, γ. Hence we may give the time axis whatever direction we choose toward the upper half of the world, $t > 0$. Now what has the requirement of orthogonality in space to do with this perfect freedom of the time axis in an upward direction? [Minkowski 1909, p. 77]

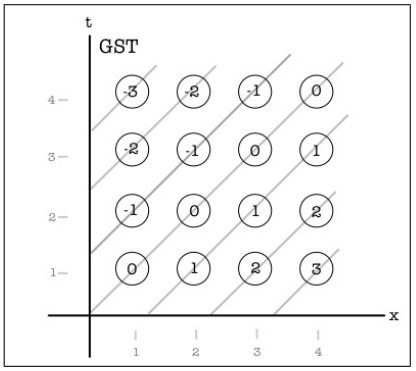


Figure 2: Values of $x' = x - \alpha t$, with $\alpha = 1$, are circled. With respect to lines of constant x' (in grey), the time axis has a different 'direction'.

Minkowski seems to suggest that, *ceteris paribus*, we should prefer coordinate transformations that do preserve that orthogonality. Thus Minkowski goes on to introduce the now-familiar line element $c^2t^2 - x^2 - y^2 - z^2 = 1$ and, after showing that the corresponding symmetry group converges to the Galilean group when $c \to \infty$, he points out that all transformations preserving this line element would also preserve the sought-after orthogonality. From this, Minkowski arrives at his stated goal of showing "how it might be possible, setting out from the accepted mechanics of the present day, along a purely mathematical line of thought, to arrive at changed ideas of space and time" [p. 75]. In short, Minkowski feels that the aforementioned features of this new line element are attractive enough that one ought, after considering the options, to employ this geometrical structure in doing spacetime physics.

> [It] looks as though the thought might have struck some mathematician, fancy-free, that after all, as a matter of fact, natural phenomena do not possess an invariance with the group G_∞ [Eqn. 1, p. 6], but rather with a group G_c [Eqn. 2, p. 8], c being finite and determinable, but in ordinary units of measure, *extremely great*. [Minkowski 1909, p. 79]

Of course, this kind of opportunity was missed for mathematicians, but Minkowski explains that this symmetry group is indeed equivalent to that of any dynamical law meeting the requirements forced upon us by empirical observation, as shown in Einstein's

work [1905*b*]. He goes on to explain that if material objects behave in accord with this geometrical structure, experiments like Michelson and Morley's would necessarily produce a null result. In fact, he emphasizes that his conception of space and time can be used to get Lorentz's contraction hypothesis more clearly than Lorentz's own work [1909, pp. 81–3]. And so it seems that, like Einstein, Minkowski agreed that a physical contraction would explain experimental results like those of Michelson and Morley. But just as Einstein preferred other foundations, Minkowski preferred postulating a non-Euclidean spacetime structure instead of relying upon an allegedly *ad hoc* contraction hypothesis.[30]

The Geometrical Approach to Relativity

Minkowski's work paved the way for what has come to be known as the 'geometrical approach' to relativity. In short, the geometrical approach puts emphasis and priority on geometrical structure in accounting for the kinematical changes entailed by Einstein's two postulates. Thus relativistic effects, inertial motion, and even the Lorentz covariance of dynamical laws are all allegedly accounted for by the world's Minkowskian geometrical structure. John Winnie, among others, promotes something along these lines.

> The special theory of relativity forced a radical revision of classical views about the causal structure of the world. [...] Foundational studies, such as that of Reichenbach [1969], soon revealed that this departure from classical causality in the special theory is intimately related to its most dramatic consequences: the relativity of simultaneity, time dilation, and length contraction. By now it has become clear that these kinematical effects [the relativity of simultaneity, time dilation, and length contraction] are best seen as consequences of the geometrical structure of Minkowski space-time, which in turn incorporates a nonclassical theory of causal structure. [Winnie 1977, pp. 134–5]

Over time, this presentation of special relativity has become the orthodox one. As Brown has pointed out [2005*b*, pp. 132ff], canonical texts like Michael Friedman's [1983] introduce special relativity in a similar way—as if spacetime structure is explanatorily prior to dynamics in accounting for the kinematics of special relativity. Consider the

[30] But see footnote 14 on page 10.

following excerpt from Friedman's text, *Foundations of Space-Time Theories*.[31]

> According to the present point of view, then, the basic or primitive elements of our theories are of two kinds: space-time and its geometrical structure; and matter fields—distributions of mass, charge, and so on—which represent the physical processes and events occurrent within space-time. Our theories seek to explain and predict the properties of material processes and events by relating them to the geometrical structure within which they are "contained." This point of view contrasts with standard philosophical formulations, such as Reichenbach's [85], which characteristically take more observational entities—reference frames, light rays, particle trajectories, material rods and clocks—as primitive and attempt to define geometrical structure in terms of the behavior of such relatively observational entities. In the present treatment we explicitly take the more abstract geometrical entities as primitive and define the more observational entities in terms of them. [Friedman 1983, p. 32]

True to form, Friedman's introduction to special relativity begins with the details of Minkowski geometry [p. 125], and in his discussion of the reality of spacetime, he does the same [pp. 253ff].[32] In Brown's review of Friedman's work, Friedman is described as taking this assumption of Minkowski spacetime to amount to what Einstein called a constructive approach to special relativity [2005*b*, p. 132]. In the next section, it will be argued that the question of whether or not Minkowski spacetime can play a role in a constructive theory has caused more confusion than clarification in the recent literature. But either way, it should be stressed that the idea underlying the geometrical approach is indeed some sort of an explanatory claim. Friedman, for instance, takes Minkowski geometry to provide a 'unificationist' explanation of relativistic phenomena and inertial motion [1983, Ch. VI.3–VI.4, esp. fn. 14].[33]

Another person whom Brown presents as an advocate of the geometrical approach is Graham Nerlich, whose geometrical explanation of inertial motion he quotes disapprov-

[31] Following on from footnote 26 (p. 16), the following quotation makes it clear that Friedman does take the position of spacetime substantivalism.

[32] Friedman gives a similar introduction to special relativity elsewhere [2002]. Janssen [2009] describes this approach as "the church" of $\langle M, O_i \rangle$, and cites Earman [1989] as another one of its promoters.

[33] Unificationist explanations are introduced in more detail in §2.2. In short, a unificationist explanation is one in which some connection is established between features of the world that were previously thought to be entirely independent.

ingly.

> [W]ithout the affine structure there is nothing to determine how the [free] particle trajectory should lie. It has no antennae to tell it where other objects are, even if there were other objects ... It *is because space-time has a certain shape that world lines lie as they do.*
>
> [Brown 2005*b*, p. 24, quoting Nerlich [1976]]

At this point, it should be acknowledged that in the second edition of Nerlich's book [1994*a*], published before Brown's work [2005*b*] or his and Pooley's collaboration [2006], this quotation does not appear. What's more, Nerlich has also taken care to explain elsewhere [1979] that his geometrical explanation does not have the causal tone that Brown and Pooley give it, and has gone so far as to argue in print that the authors misunderstood his position [2010, §5; 2013, Ch. 8]. Still, whether or not the account is causal, Nerlich is certainly a strong advocate of the geometrical approach, and some of his work in favor of geometrical explanations is explored further in Chapter 2.[34]

While Friedman and Nerlich appear to be the targets of Brown's criticisms, which are outlined in more detail below, other advocates of a geometrical approach to special relativity have come to the fore in recent years. One party that should be mentioned is Dorato and Felline [2010], who take Minkowski spacetime structure to provide what they call, following Hughes [1989, 1993], a 'structural' explanation of relativistic phenomena and inertial motion. While this is not the place for a full introduction of structural explanations, suffice it to say that the authors account for the aforementioned explananda by locating them in Minkowski spacetime *qua* a mathematical model, so that "causation, mechanical models or dynamical forces are never called into play" [p. 197]. Another such party is Balashov and Janssen [2003], who defend a geometrical explanation somewhat similar to Friedman's unificationist one. Their work, and Janssen's subsequent independent work, is outlined in more detail below. Finally, a geometrical approach also appears in Tim Maudlin's recent introduction to the philosophy of space and time [2012]. Although Maudlin does not engage directly with the literature outlined in this chapter,

[34] Bradford Skow has argued that Nerlich's geometrical explanation is indeed causal [2013, §3].

his introduction to special relativity has an emphasis on the geometrical structure of Minkowski spacetime that might be compared to that of Friedman, mentioned above.[35]

The Arrow of Explanation

Despite the diversity in the specific forms of explanation suggested by Friedman, Nerlich, and others, a common element to those approaches is the priority of Minkowski geometrical structure in accounting for Lorentz covariance. Indeed, many proponents of a geometrical approach to special relativity actually acknowledge that Lorentz covariance does in fact account for relativistic phenomena, but go on to argue that in turn, Lorentz covariance is accounted for by Minkowski spacetime structure. That is, whereas in 1905 Einstein took a principle-theory approach in deriving the 'restrictive principle' that all laws must be Lorentz covariant, most advocates of a geometrical approach begin with the assumption of Minkowskian geometrical structure, and then argue that dynamical laws' symmetries reflect, or perhaps 'respect', that structure. This is the 'arrow of explanation' from Minkowski geometry to Lorentz covariance, and is the explanatory claim that takes center stage for most of this thesis.

This idea comes to the surface in some of the aforementioned works. Consider the following quotation by Friedman.[36]

> Finally, what is the symmetry group of relativistic electrodynamics? [...] The absolute objects of our theory are just D [the affine connection] and g [the metric]; the fixed background space-time structure of our theory is just Minkowski spacetime. It follows that the symmetry group of relativistic electrodynamics is just the Lorentz group. [Friedman 1983, p. 149]

Granted, Friedman refers here to the theory's 'symmetry group' rather than its 'covariance group', but this amounts to a mere difference in terminology. For as Friedman explains earlier in the same work, he uses the former term to refer to the covariance group of the theory's standard formulation—i.e., the group under which the standard formulation of

[35] See especially Maudlin's discussion of Lorentz-covariance and length contraction [2012, pp. 116ff.].
[36] As mentioned in footnote 10 (p. 9), what Friedman means by the 'Lorentz group' is being called the 'Poincaré group' in this thesis.

the theory preserves its form.[37] And so Friedman maintains that the arrow of explanation runs from geometry to dynamics.

Something similar can be seen in the work of Nerlich, whose outline of the geometrical approach presents it as involving an ontology of four-dimensional objects in some four-dimensional spacetime [1994*b*, p. 63]. And although his description focuses mainly on the ontology, it is also clear that he has this explanatory priority in mind.

> It is, in fact, the metric structure of spacetime which bestows on the Lorentz coordinates (and their connecting group) the special symmetries which make possible the simple forms of physical law. [Nerlich 1994*b*, p. 66]

For the physical laws to take their 'simple form' with respect to coordinatizations that are related by Lorentz transformations is for the laws to be Lorentz covariant. But instead of viewing that characteristic of dynamical laws as a given, Nerlich takes it to be accounted for by the Minkowski metrical structure.[38] Indeed, Nerlich goes on to defend this view much more explicitly in later work [2013], which is introduced in more detail below (Ch. 2).

Finally, this arrow of explanation has been advocated more recently by Balashov and Janssen [2003], whose work comes up again in the following sections.[39]

> [Does] the Minkowskian nature of space-time explain why the forces holding a rod together are Lorentz invariant or the other way around? Our intuition is that the geometrical structure of space(-time) is the *explanans* here and the invariance of the forces the *explanandum*. To switch things around, our intuition tells us, is putting the cart before the horse.
>
> [Balashov and Janssen 2003, pp. 340–1]

And quite similar to Friedman's outline of a unificationist form of geometrical explanation, Balashov and Janssen defend a 'common origin inference', according to which Minkowski spacetime accounts for the coincidence that all natural laws share the same symmetry

[37] Note that Friedman refers to the 'symmetry group' and 'invariance group' of a theory interchangeably [1983, p. 56, pp. 60–1]. The former term (and its definition) he gets from Anderson [1976, pp. 84–8].

[38] Craig [2001, p. 80, fn. 22] makes a similar observation. In personal correspondence, Nerlich has confirmed this interpretation of the quotation.

[39] The terms 'explanans' and 'explanandum' are defined on page §2.1, below.

group [pp. 341–2]. The inference is made primarily by assuming the validity of Earman's well-known symmetry principles.

> SP1: Any dynamical symmetry of [theory] T is a space-time symmetry of T
> SP2: Any space-time symmetry of [theory] T is a dynamical symmetry of T
> [Earman 1989, p. 46]

As with Friedman's terminology above, Earman has a particular definition in mind for dynamical symmetries and spacetime symmetries.[40] In short, a dynamical symmetry is an element of the covariance group of the dynamical laws, and a spacetime symmetry is a symmetry of the geometrical structure of spacetime. With those definitions, it is clear that if the two principles hold, and if Minkowskian geometrical structure is assumed from the start, then the Lorentz covariance of dynamical laws is entailed [pp. 341–2]. Of course, Earman's principles could also be applied in an argument from the Lorentz covariance of dynamical laws to the Minkowski geometry of spacetime, but the authors much prefer the assumption of Minkowski spacetime structure over assuming the independent Lorentz covariance of all dynamical laws.

The idea of a common origin inference was introduced in earlier work by Janssen [2002*a*]. And in his later work [2009], Janssen takes them further. There, Janssen takes a primarily historical approach; he spends considerable time summarizing the historical development of explanations for three relativistic phenomena, each of which played a significant role in the early reception of special relativity: the Fresnel drag effect; the velocity dependence of electron mass; and the torques on a moving capacitor. His main point in doing so is to argue that these effects, along with time dilation and length contraction, were eventually given—and should continue to be given—*kinematical*

[40] Earman defines spacetime symmetries to be those diffeomorphisms that map spacetime onto itself while preserving its absolute geometrical structures. Or, in the terminology introduced in footnote 32 (p. 20), a spacetime symmetry of a spacetime model $\langle M, A_1, A_2, ..., P_1, P_2, ...\rangle$ is "a mapping that leaves all of the A_is invariant, i.e., a diffeomorphism ψ that maps M onto M in a way that $\psi * A_i = A_i$ for all i", where the A_i are the absolute geometric-object fields representing fixed spacetime structure [Earman 1989, p. 49]. As for dynamical symmetries, Earman defines them to be diffeomorphisms ϕ that map M onto M by acting on the P_i, which represent the physical contents of spacetime, in a way that the dynamical laws do not distinguish [pp. 45–6]. In other words, the dynamical symmetries map solutions to solutions, and non-solutions to non-solutions. Note that the covariance group of an equation can also be defined as the group under which the equation itself preserves its form. (See page 9, above.)

explanations rather than *dynamical* ones. And because Janssen believes that kinematical effects do not stand in need of any dynamical explanation, he describes his disagreement with Brown and Pooley as "ultimately about how to draw the line between kinematics and dynamics in special relativity" [p. 27]. Janssen traces this distinction in several historical examples, but the heart of his message is the same: kinematical phenomena are accounted for by features of the dynamical laws that are to be expected after assuming Minkowski geometrical structure [pp. 48–50].

1.3 The Dynamical Approach to Special Relativity

In recent years, Harvey Brown and Oliver Pooley have argued against the geometrical approach to special relativity, and in favor of what I will call the 'dynamical approach'. In *Physical Relativity*, Brown gives some prehistory to the dynamical approach by outlining the work of various individuals whose interpretation of relativity was more dynamical than geometrical [2005*b*, Ch. 7]. However, there is less attention given to the more immediate history of the dynamical approach. Some of that development is given below, focusing primarily upon the works of John Bell and William Lane Craig, from which the dynamical approach is then distinguished. Finally, Brown and Pooley's collaborations are introduced together, alongside corresponding parts of Brown's book.

Bell's Lorentzian Pedagogy

In Brown's outline of unconventional voices on special relativity, a key player is John Bell, who is included primarily as an advocate of a 'Lorentzian pedagogy' for special relativity.[41] In his essay *How to Teach Special Relativity* [1987], Bell offers an approach to special relativity that he contrasts with the "usual approach". Whereas the latter allegedly emphasizes "the radical break with more primitive notions of space and time", Bell's

[41] [Brown 2005*b*, §1.4]. Brown has repeatedly pointed out that Bell refers to the 'Lorentz contraction' and 'FitzGerald contraction' interchangeably, and refers to what should be called the 'FitzGerald pedagogy' as the 'Lorentz pedagogy' [Brown and Pooley 2001, 2006; Brown 2003; Brown 2005*b*, pp. 6–7]. For whereas Bell's approach involves no interaction between matter and the ether, Lorentz continued to pursue a mechanical account of length contraction along those lines. Still, I follow suit and refer to Bell's approach as the 'Lorentzian pedagogy'.

introduction is claimed to "emphasize the continuity with earlier ideas", and to thereby have certain pedagogical advantages [p. 67]. He admits that his approach is essentially a longer, more arduous route to the same results, but his point is simply that "[the] longer road sometimes gives more familiarity with the country" [p. 77]. In another context, Bell explains:

> If you are, for example, quite convinced of the second law of thermodynamics, of the increase of entropy, there are many things that you can get directly from the second law which are very difficult to get directly from a detailed study of the kinetic theory of gases, but you have no excuse for not looking at the kinetic theory of gases to see how the increase of entropy actually comes about. In the same way, although Einstein's theory of special relativity would lead you to expect the FitzGerald contraction, you are not excused from seeing how the detailed dynamics of the system also leads to the FitzGerald contraction. [Bell 1992, p. 34]

Now, given that Einstein himself was motivated by thermodynamics in developing his special 'principle-theory' of relativity, it should be clear that Bell is offering something along the lines of a constructive version of special relativity, and that it is for this reason that he refers to his work as defending the 'Lorentzian pedagogy'.[42]

Bell illustrates this pedagogy by describing the dynamics of an electron in orbit around a nucleus, ignoring any back-effect of the electron on the nucleus itself.[43] Using the relativistic form of Maxwell's equations, Bell programs a computer to integrate the equation of motion for the electron (Lorentz's force law), while assuming a relativistic equation relating the electron's velocity and mass.[44] In doing so, Bell shows this dynamics to have the following three features.

1. When the system is set into motion and observed from the frame in which it was previously at rest, the orbit of the electron will have contracted in the direction of its motion by a specific amount.

[42] But see footnote 41, above.

[43] Bell acknowledged that a full description of the dynamics would have to include a quantum theory of stable matter. Let us imagine, for the sake of illustration, that Bell provided this kind of detail.

[44] That is, in solving $\frac{d\mathbf{p}}{dt} = -e(\mathbf{E} + c^{-1}\dot{\mathbf{r}}_e \times \mathbf{B})$, Bell uses $\dot{\mathbf{r}}_e = \mathbf{p}/\sqrt{m^2 + \mathbf{p}^2 c^{-2}}$ instead of $\dot{\mathbf{r}}_e = \mathbf{p}/m$. This is effectively Bell's way of exercising Einstein's restrictive principle that the dynamical laws be Lorentz covariant (see p. 15, above).

2. When the system is set into motion and observed from the frame in which it was previously at rest, the period of its orbit is dilated by a specific amount.

3. There exist other (primed) coordinates which can be used to describe the moving system in the same way as the original coordinates were used to describe the system at rest, and which are also the same coordinates one obtains by performing a Lorentz transformation from the 'rest' frame to the frame comoving with the system.[45]

Thus Bell's pedagogy gives a dynamical account of length contraction and time dilation. Indeed, his example is similar in spirit to that of Heaviside [1888], whose work first inspired FitzGerald to make the contraction hypothesis (see §1.1). His take-away point is that it is possible to do one's physics from any single frame, 'as if' that frame were the one true frame. But another lesson not to be overlooked is that this can be done without fully adopting Lorentz's metaphysics of a preferred frame.

> The facts of physics do not oblige us to accept one philosophy rather than the other. And we need not accept Lorentz's philosophy to accept a Lorentzian pedagogy. Its special merit is to drive home the lesson that the laws of physics in any *one* reference frame account for all physical phenomena, including the observations of moving observers. [Bell 1987, p. 77]

Craig's Neo-Lorentzian Approach

While Bell would argue that Lorentz's commitment to a preferred frame is not a necessary step in adopting a Lorentzian pedagogy, William Lane Craig argues that one ought to acknowledge a preferred frame of reference. Alongside Bell's work, Craig's neo-Lorentzian approach provides an important contrast to the dynamical approach, the latter being sometimes mistaken for the former.

In a three-volume work [2000*a*; 2000*b*; 2001], Craig outlines a "coherent doctrine of divine eternity and God's relationship to time" [2001, ix]. Its third volume [2001] includes

[45] Bell suggests that by this third feature, he has given a dynamical derivation of Lorentz covariance for the dynamics. But Lorentz covariance implies that the above feature would hold between any two inertial frames for any solution to the dynamics, whereas Bell has only shown this for one particular solution. This can be overlooked for present purposes.

a proposal for a neo-Lorentzian approach to special relativity.[46] In it, Craig describes the task of interpreting special relativity as a matter of choosing between rival theories. One of those 'theories' has already been introduced as the geometrical approach, which Craig calls the 'space-time interpretation'.[47] But to be clear, Craig's space-time interpretation has a fundamental ontology that includes substantival spacetime, meaning that in Craig's view, the geometrical approach accounts for relativistic phenomena and the Lorentz covariance of dynamical laws by an appeal to the geometrical structure of Minkowski spacetime *the entity*.[48] Craig sings the praises of the geometrical approach, arguing that it "has greater explanatory power than [Einstein's] interpretation" because "phenomena which are simply brute facts according to [Einstein's] interpretation can be explained within [it], thus increasing our understanding of the physical world" [p. 102]. But this actually marks both the beginning and the end of Craig's praise for the geometrical approach. Although he agrees with its potential explanatory power, his discomfort with the geometrical approach comes from his adherence to an A-theory of time, in favor of which he argues in the earlier volumes of his project, and which he believes is incompatible with spacetime substantivalism.[49] Thus Craig sets the goal of seeking out "a physical interpretation of the SR formalism predicated upon a coherent metaphysic which permits us to affirm both a theory of tensed time and a unified view of reality" [p. 104]. A few chapters later, Craig invokes the existence of God and an A-theory of time in an argument for the existence of a preferred frame of reference [p. 173], which finally leads him to propose another so-called theory: a 'neo-Lorentzian' approach to special relativity

[46] More recently, Craig has published another article [2008], which includes several excerpts from his earlier work [2001]. Most of the following quotations can be found repeated there.

[47] Craig cites Nerlich for these names [Nerlich 1994*b*, p. 63; 1998, pp. 128–9]

[48] As mentioned at the start of §1.2, this is not the case for some advocates of the geometrical approach. Craig takes this ontology to have followed from Minkowski's work of 1909, though another interpretation of Minkowski's work is defended in §3.2, below.

[49] The A-Theory of time is a tensed theory of time, according to which events are described as being past, present, or future. Events are ordered by their having tensed, monadic properties such as 'being two days in the future', or 'being one day in the past'. It is contrasted with the 'tenseless' B-Theory of time, according to which events are described only relatively. According to the B-Theory, events are ordered by their having dyadic properties such as 'being two days later than', or 'being one day earlier than'. The terms are due to McTaggart [1908].

[p. 178].[50]

For present purposes, we can ignore the details of Craig's neo-Lorentzian proposal, and summarize it as a three-dimensional picture with a preferred frame of reference, in which some particular, Lorentz-covariant dynamical laws govern the behavior of material objects. What is relevant here is Craig's claim that the neo-Lorentzian approach to special relativity provides the best explanation of relativistic effects. Later in the same work he claims, "As a constructive theory, the neo-Lorentzian approach promises to enrich our understanding of the causal structure of the world in a way that Einstein's cannot" [p. 182]. Here again, Craig grants that the geometrical approach "does explain why these relativistic phenomena occur", but calls into question "whether any such metaphysical reality as spacetime actually exists" in the first place [p. 182]. In the end, he claims that "commitment to an A-Theory of time, and thus to the objectivity of tense and temporal becoming, is itself sufficient for the rejection of spacetime realism" [p. 192]. Indeed, in 2008, Craig even states that his 2001 argument can be freed from its theistic underpinnings:

> For the yeoman's work in this argument is done, not [by] the assumption of theism, but rather by the assumption of an A-Theory of time. One may plausibly argue directly, I think, that if an A-Theory of time is correct, then a Lorentzian interpretation of SR is correct. [Craig 2008, p. 22]

While Craig's views remain unorthodox in the philosophy of special relativity, they had a significant role in sparking recent controversies between the geometrical and dynamical approaches to relativity. For it was in response to Craig's work that Balashov and Janssen wrote their defense of the geometrical approach to relativity, introduced above [2003]. And it was their defense of the geometrical approach to which Brown and Pooley responded in the collaboration [2006] that gave the dynamical approach the form

[50] Craig calls the third theory 'the relativity interpretation'. It involves an ontology that Craig believes Einstein to have advocated in his 1905 derivation of the Lorentz transformations: a three-dimensional world in which all inertial frames of reference have been relativized to an equal status, resulting in a frame-dependent ontology for all material objects. Although Craig acknowledges a few other authors' arguments that Einstein's did not intend to take *any* position on the ontology of spacetime, Craig stands firm that this was the picture that Einstein had in mind [2001, pp. 24–5].

in which it appears in Brown's book [2005*b*, Ch. 8].

The Dynamical Approach

In brief, the dynamical approach argues for the priority of dynamics over geometry in the context of explanation: relativistic kinematical phenomena like length contraction and time dilation, along with inertial motion and even Minkowskian geometrical structure, are to be explained by the fact that the dynamical laws governing the behavior of material objects are Lorentz-covariant.[51] Brown's efforts along these lines can be traced from some of his earlier work [1992; 1993; 1995; 1997; 2003] into his collaborations with Oliver Pooley [2001; 2006], and another with Chris Timpson [2006]. With the exception of the latter, these culminate in his book *Physical Relativity* [2005*b*].

In the first of their collaborations, Brown and Pooley promoted what they called a "truncated Lorentzian pedagogy" [2001, p. 261], which was a variation of the 'Lorentzian pedagogy' promoted by Bell. Like Bell, they stress that one needn't adopt a preferred frame to adopt a Lorentzian pedagogy, but they also stress that one needn't focus on the specific dynamics of either Lorentz's theory of electrons or Bell's simplified atomic model.[52] That is, the authors point out that an important lesson of Bell's paper is that an explanation of relativistic phenomena needn't appeal to the full details of a system's dynamics, but can (and should) instead appeal to particular characteristics of a system's dynamics—namely, their Lorentz covariance.[53]

> [...] Bell did not seem to believe that articulation of a *complete* dynamical treatment of this kind was a necessary part of the Lorentzian pedagogy. In order to predict, on dynamical grounds, length contraction for moving rods and time dilation for moving clocks, Bell recognized that one need not

[51] See *Physical Relativity* [2005*b*, pp. vii–viii, pp. 24–5, Ch. 8]. Reviews of *Physical Relativity* by Huggett [2009] and Skow [2006] discuss the way that these explanatory issues all fit together, and are each cited below. Other reviews not cited here include those by Dorato [2007] and Martínez [2007]. Finally, Pooley's own perspective on the dynamical approach can be found in his recent overview of the substantivalism–relationalism debate [2013], and is discussed at length in Chapter 5, below.

[52] But of course Bell did not actually provide a complete dynamics. See footnote 43 on page 26. Note also that, as suggested in the quotation below, Bell did not believe that the full details of the dynamics should be appealed to in an explanation of relativistic effects.

[53] See also the comments by Brown [2005*b*, p. vii] or Brown and Pooley [2006, p. 82].

> know exactly how many distinct forces are at work, nor have access to the detailed dynamics of all of these interactions or the detailed micro-structure of individual rods and clocks. It is enough, said Bell, to assume Lorentz covariance of the complete dynamics – known or otherwise – involved in the cohesion of matter. We call this the *truncated* Lorentzian pedagogy.
>
> [Brown and Pooley 2001, p. 261]

And although the name didn't stick, this same perspective was defended in their second collaboration, in which the authors replied to Balashov and Janssen's work, introduced above.

> [In] many contexts, perhaps in most contexts, one should not appeal to the *details* of the dynamics governing the microstructure of bodies exemplifying relativistic effects when one is giving a constructive explanation of them. *Granted that there are stable bodies*, it is sufficient for these bodies to undergo Lorentz contraction that the laws (whatever they are) that govern the behaviour of their microphysical constituents are Lorentz covariant. It is *the fact that the laws are Lorentz covariant*, one might say, that explains why the bodies Lorentz contract. To appeal to any further details of the laws that govern the cohesion of these bodies would be a mistake.
>
> [Brown and Pooley 2006, p. 82]

Now, in this later work, the authors go so far as to acknowledge certain contexts in which Minkowskian geometrical structure is indeed explanatory of relativistic phenomena. But still, they stress that even in those contexts, such explanations are not what Einstein would have called 'constructive', *pace* Balashov and Janssen [pp. 78–9]. Their reasons for doing so are outlined in Chapter 2, below. And whereas the Lorentzian pedagogy had previously focused primarily upon relativistic phenomena, it becomes clear in their 2006 collaboration that the dynamical approach also takes a side in the question of the arrow of explanation between Lorentz covariance and Minkowski geometry. In response to Balashov and Janssen's presentation of the geometrical approach, Brown and Pooley point out that no account has been given as to how it is that spatiotemporal geometry can influence the symmetry group of dynamical laws.

> Craig's neo-Lorentzian interpretation is precisely an example of a theory in which the symmetries of space-time structure are not reflected in the symme-

> tries of the laws governing matter. [...] This shows that, as matter of logic alone, if one postulates space-time structure as a self-standing, autonomous element in one's theory, it need have no constraining role on the form of the laws governing the rest of content of the theory's models. So how is its influence on these laws supposed to work? Unless this question is answered, space-time's Minkowskian structure cannot be taken to explain the Lorentz covariance of the dynamical laws.
>
> [Brown and Pooley 2006, p. 84; *cf.* Brown 2005*b*, §8.3.2]

Of course, this is not to say that Brown and Pooley advocate a position that would deny Earman's symmetry principles.[54] For according to the dynamical approach, the spacetime symmetry group and dynamical symmetry group do indeed match, but the arrow of explanation runs in the opposite direction: Minkowski spacetime structure is accounted for by the symmetries of Lorentz-covariant dynamical laws.

> [Minkowski spacetime] is definable in terms of the Lorentz covariance of the fundamental physical laws. [...] In the context of SR, rods and clocks *are* surveying devices for a four-dimensional geometric structure. But this is a structure defined in terms of the symmetries of the dynamical laws. If matter and its interactions are removed from the picture, Minkowski spacetime is not left behind. [Brown and Pooley 2001, p. 269]

And once again, the same point is argued in their 2006 reply to Balashov and Janssen.

> It is the Lorentz covariance of the laws that underwrites the fact that the geometry of space-time is Minkowskian. It is for this reason that we can rule out the sort of mismatch between space-time symmetries and dynamical symmetries that are a feature of Craig's interpretation, and that so trouble Balashov and Janssen. [Brown and Pooley 2006, p. 84]

Finally, it should be noted that by the authors' use of the term 'Minkowski spacetime' in the penultimate quotation above, the dynamical approach comes across as a relationalist position, according to which (Minkowski) spacetime does not exist primitively. More is

[54] That said, the quotation above does show that Earman's SP1, introduced above on page 24, is not a logical truth. See also the related discussion in footnote 7 on page 98. More recently, Pooley has acknowledged that although the symmetry group of the dynamical laws governing matter may be larger than that of any "self-standing, autonomous" spacetime structure, the former should at least include the latter [2013, p. 571; 2014, §6.2].

said in §2.3 about where the dynamical approach fits in the substantivalist–relationalist spectrum. But as suggested by the comments at the start of §1.2, the explanatory claims being outlined in this section do not hinge upon whether or not substantival spacetime features in the fundamental ontology. That is, although some authors like Janssen [2009] take Brown to defend relationalism in *Physical Relativity* [2005*b*], Brown makes it clear that "the dynamical version of relativity theory is a separate issue and can be justified on much wider grounds, having essentially to do with good conceptual house-keeping" [2005*b*, p. ix]. Elsewhere, Brown and Pooley also maintain that their arguments against the constructive nature of Balashov and Janssen's geometrical explanation applies to substantivalists and relationalists alike.[55]

> What needs to be stressed is that this conclusion is appropriate not only for those who adopt an eliminative relationalist stance towards the ontology of space-time, and not only in the context of theories with fixed, absolute space-time structure. As we argued in Brown and Pooley [2001], even when one's ontology *includes* substantival space-time structure, the symmetries of the laws governing material systems are still crucial in such structure gaining operational chronogeometric significance. [Brown and Pooley 2006, p. 86]

Summary

In sum, we have seen that in distinction to Einstein's principle-theory approach of 1905, two other approaches to special relativity offer rival accounts of relativistic phenomena and either Minkowski geometry or the Lorentz covariance of dynamical laws. But to be clear, although those approaches take opposing sides regarding the arrow of explanation between Minkowski geometry and Lorentz covariance, their differences are subtler when it comes to relativistic effects. For while relativistic effects are accounted for by Lorentz covariance in the dynamical approach, many advocates of the geometrical approach would grant this explanation as well, but would then take one step further in arguing that the laws' Lorentz covariance is in turn accounted for by Minkowski geometry. The geometrical approach has received criticism primarily from Brown and Pooley, although others have

[55] See also footnote 31 on page 65, and the related discussion in §6.1.

voiced similar concerns.[56] The dynamical approach has received a number of responses in the literature, most of which have focused on the explanatory claims mentioned above.[57] The goal here has been to present the explanatory claims while avoiding related issues like the substantivalist–relationalist debate, so that a more critical review of the recent literature can be given in the following section.

[56] DiSalle is one example [2006], whom Brown quotes approvingly.

[57] In addition to the works cited above, see comments by Dieks [2009], Felline [2011], Frisch [2011], Lange [2011], and van Camp [2011*a*].

2 Einstein's Legacy and Constructive Relativity: A Critical Review of an Ongoing Debate

As noted above, Einstein derived the Lorentz transformations from two seemingly incompatible postulates. But although Einstein showed that reconciling those postulates would entail relativistic effects, it could be argued that he did not thereby explain relativistic effects. Indeed, Einstein himself felt this way: on several occasions Einstein claimed that 'principle-theories'—of which his 1905 work was an example—are not explanatory. As a result, some authors have since felt a need to bolster any account of relativistic effects by showing it to meet Einstein's criteria for a so-called 'constructive theory', to which he did ascribe explanatory power. The first aim of this chapter (§2.1) is to introduce Einstein's distinction in more detail, and to argue that it should not be taken as an exhaustive taxonomy of physical theories, or even as an acceptable framework for the explanatory power of various physical theories. The second aim of this chapter (§2.2) is to show that Einstein's distinction has set the stage somewhat awkwardly for an evaluation of the rival approaches to special relativity outlined in Chapter 1. That some authors seem to tacitly endorse Einstein's claim while taking different views on what counts as a 'constructive' (and therefore explanatory) theory has led to some misunderstandings in the recent literature. Only with a broader framework than Einstein's principle–constructive distinction can the explanatory power of these rival approaches be appreciated.[1]

Separate and aside from Einstein's principle–constructive distinction is the fact that Brown and Pooley's dynamical approach to relativity has a certain characteristic that lends it to be described, perhaps somewhat confusingly, as 'constructive relativity'. In particular, it is the claim that Minkowski spacetime structure can be recovered from the symmetries of the dynamical laws without assuming that structure from the start. (To be clear: this should not be taken to mean that the dynamical approach is a 'constructive theory', as Einstein uses the term.) The third aim of this chapter (§2.3) is to outline and

[1] Although some of the literature cited below also focuses on explanation of inertial phenomena, this chapter focuses primarily upon explanations of relativistic effects, as well as the arrow of explanation between Lorentz covariance and Minkowskian geometrical structure.

respond to certain criticisms of this particular aspect of the dynamical approach. Thus the focus shifts away from the issues discussed in the first two sections, and over to some more metaphysical issues related to the claim that the symmetries of the dynamical laws account for those of spacetime geometry. Still, a common theme with earlier sections of this chapter is clearing up misunderstandings or misrepresentations of the dynamical approach in the literature since Brown and Pooley's work from 2001–2006. And with regard to the criticisms raised in this third and final section, it is the collective aim of later chapters to provide an adequate response. Therefore the conclusion of this chapter is open-ended, motivating the project to be undertaken in Chapters 3–5.

2.1 On Einstein's Principle–Constructive Distinction

Though he may not have been the first to make it, Einstein is known for popularizing the distinction between so-called 'constructive theories' and 'principle-theories'. And as mentioned above, Einstein seemed to recognize explanatory power only in the former. However, work on scientific explanation over the last half-century has made clear the difficulties in developing any one 'model' of scientific explanation, or pinpointing any one kind of understanding that science provides. Even when the conversation is restricted to physics, various different examples of scientific explanation can be given. After introducing Einstein's principle–constructive distinction in more detail, a brief overview of recent work on scientific explanation is given below. Finally, it is argued that Einstein's principle–constructive distinction should be replaced by a broader, more pluralistic framework for the explanatory power of physical theories.

The Principle–Constructive Distinction

One of Einstein's best-known presentations of the principle–constructive distinction was published in the *London Times* [1919].

> We can distinguish various kinds of theories in physics. Most of them are constructive. They attempt to build up a picture of the more complex phenomena out of the materials of a relatively simple formal scheme from which

> they start out. Thus the kinetic theory of gases seeks to reduce mechanical, thermal, and diffusional processes to movements of molecules—i.e. to build them up out of the hypothesis of molecular motion. When we say that we have succeeded in understanding a group of natural processes, we invariably mean that a constructive theory has been found which covers the processes in question.
>
> Along with this most important class of theories there exists a second, which I will call "principle-theories." These employ the analytic, not the synthetic, method. The elements which form their basis and starting-point are not hypothetically constructed but empirically discovered ones, general characteristics of natural processes, principles that give rise to mathematically formulated criteria which the separate processes or the theoretical representations of them have to satisfy. Thus the science of thermodynamics seeks by analytical means to deduce necessary conditions, which separate events have to satisfy, from the universally experienced fact that perpetual motion is impossible.
>
> The advantages of the constructive theory are completeness, adaptability, and clearness, those of the principle theory are logical perfection and security of the foundations.
>
> The theory of relativity belongs to the latter class.
>
> [Einstein 1919, p. 228]

By his description, Einstein's taxonomy of theories includes (at least) constructive theories and principle-theories.[2] And with regard to the claim that his 1905 derivation of the Lorentz transformations is a principle-theory, it would be hard to disagree. As explained in §1.1, Einstein took two empirically supported regularities and proposed their universal truth as a starting point, but gave no information about the specific mechanics or dynamics of rods or clocks. Along these lines, an anecdote reported by Brown is fitting.

> The one and only meeting between Einstein and the great French polymath took place at the Solvay Conference in Brussels in 1911, a year before Poincaré's death. Maurice de Broglie later recalled: 'I remember one day at Brussels, while Einstein was explaining his ideas, Poincaré asked him, "what mechanics are you using in your reasoning?" Einstein answered: "No

[2] Einstein gives a very similar introduction to the principle–constructive distinction elsewhere [1948]. Einstein also gives further comments on the role of phenomenological principles in physical theories elsewhere [1914; 1936].

> mechanics" which appeared to surprise his interlocutor.
>
> [Brown 2005*b*, p. 147, quoting Galison [2003]]

And indeed, it was for this very reason that Lorentz was less than satisfied with aspects of Einstein's work. In comparing it with his own, Lorentz observed,

> [...] the chief difference [is] that Einstein simply postulates what we have deduced, with some difficulty and not altogether satisfactorily, from the fundamental equations of the electromagnetic field. [Lorentz 1916, p. 230]

Granted, Lorentz deduced these results only in the case of electromagnetism, whereas Einstein derived Lorentz covariance more generally. But of course Lorentz is right to pick up on the difference between his and Einstein's approach to Lorentz covariance, which is precisely the difference between a principle-theory and constructive theory.

But while it may be clear and uncontroversial that Einstein's 1905 derivation of the Lorentz transformations fits his description of a principle-theory, other aspects of Einstein's distinction call for clarification. Brown and Timpson, for example, remind us not to take Einstein's terminology to suggest that what makes his 1905 work a 'principle-theory' was that fact that it involved principles.

> [Einstein's distinction] is clearly not categorical: all theories have principles, it is just that some are more phenomenological than others. Thermodynamics and statistical mechanics are on opposite ends of a spectrum of possible theories, and there are indeed respectable theories [...] which lie somewhere in between. [Brown and Timpson 2006, p. 32]

What's more, even if 'principle-theories' are defined as those theories founded upon purely phenomenological principles, DiSalle has pointed out that it is not only such theories that provide "mathematically formulated criteria" to be satisfied by particular processes or their "theoretical representations". In fact, DiSalle points out that providing some such constraint is common to all classical spacetime theories, including Newton's work.[3]

> Newtonian space-time, as characterized by the Galilean symmetry group, does indeed define a set of general constraints that all physical laws must

[3]Indeed, Newton himself made a similar distinction between the framework of principles governing all forces and the theories describing specific interactions [Flores 1999, p. 123].

> obey; it thus defines a general framework for inquiry, that is, a framework within which we can investigate forces of nature such as gravity.
>
> [DiSalle 2006, p. 119]

DiSalle describes the provision of such constraints as one of the 'functional' aspects of Einstein's distinction. He draws the lesson, and the terminology, from Flores [1999], who offers a very helpful exposition of Einstein's distinction by outlining its ontological, epistemic, and functional aspects. Regarding ontology, principle-theories postulate general principles bereft of any metaphysical commitments, while constructive theories postulate the existence of particular entities; regarding epistemology, the foundations of a principle-theory are discovered by empirical means, whereas the hypothetical elements of a constructive theory needn't have been; and regarding function, principle-theories provide criteria which must be satisfied by constructive theories, while the latter "build up a picture of the more complex phenomena" out of the hypothetical entities. Flores points out that with regard to the ontological aspect of Einstein's distinction, theories can be found that do not fit well into either category [1999, §3]. And indeed, this was Brown and Timpson's point, above. Instead, Flores argues that a clearer division of theories can be made by focusing on the functional aspect of Einstein's distinction. Along these lines, Flores suggests new terminology: let 'framework theories' be those that "impose restrictions on other *theories*" and "provide the scaffolding upon which other theories are built"; and let 'interaction theories' be those that "describe specific physical processes *within* the constraints imposed by the principles (or one of the consequences of) a framework theory" [1999, p. 129].

Flores's terminology will come in handy below, where the primary focus is another functional aspect of Einstein's distinction: the provision of explanations for laws and/or particular phenomena. That is, for present purposes, the take-home message of Einstein's distinction will be that principle-theories serve as a restrictive guide in the development of constructive theories, which alone provide us with a deeper understanding about how and why the world behaves in the way that it does. But given the wealth of literature on

scientific explanation, and the various types of understanding that different models of explanation provide, Einstein's distinction could only be taken as a fully comprehensive outline of physical theories' explanatory power if either principle-theories were also granted some form of explanatory power, or if his description of constructive theories were taken very loosely and inclusively. To illustrate just how broad and complex is the topic of scientific explanation, and to draw out the differences in the type of explanation provided by both principle-theories and constructive theories, a brief review is in order.[4]

On Models of Scientific Explanation

Debates over what constitutes a scientific explanation are longstanding, but contemporary work on the topic begins with the Deductive–Nomological model, promoted most prominently by Hempel and Oppenheim [1948].[5] It is from the Deductive–Nomological model that we get the terms *explanans* (that which does the explaining) and *explanandum* (that which is explained). The Deductive–Nomological model gets its name by the requirement that the explanans be true and include at least one law—i.e. that it be 'nomological'—and that the explanandum be a logical, 'deductive' consequence of the explanans. Given the key role of laws in this model of explanation, it is sometimes referred to as the 'covering law' model of explanation: the explanandum is accounted for by showing it to be subsumed by some law.[6] For example, Hempel and Oppenheim show that the initial drop and subsequent rise of mercury in a thermometer that has been immersed in hot water can be explained by pointing out that the water's heat first brings about an expansion in the glass thermometer, allowing the mercury to drop slightly, and only later brings about a much greater expansion in the mercury itself, resulting in its rise. They then go on to explain the role of laws in the explanation.

[4] The primary source for what follows is Woodward's introduction to scientific explanation [2010].

[5] A summary of responses to the Deductive–Nomological model, as in the outline to follow, is provided by Salmon [1989]. Salmon also provides a more concise version of this history elsewhere [1998, Ch. 19].

[6] This raises questions about what should count as a law. For instance, whereas the Deductive–Nomological model makes use of deterministic laws, Hempel also argued that statistical laws can be put to use in one of two variations of the Deductive–Nomological model: the Deductive–Statistical and Inductive–Statistical models [1965]. These and other issues regarding lawhood are discussed by Woodward [2010, pp. 8–10].

> This account consists of statements of two kinds. Those of the first kind indicate certain conditions which are realized prior to, or at the same time as, the phenomenon to be explained; we shall refer to them briefly as antecedent conditions. [...] The statements of the second kind express certain general laws; in our case, these include the laws of the thermic expansion of mercury and of glass, and a statement about the small thermic conductivity of glass. The two sets of statements, if adequately and completely formulated, explain the phenomenon under consideration: They entail the consequence that the mercury will first drop, then rise. Thus, the event under discussion is explained by subsuming it under general laws, i.e., by showing that it occurred in accordance with those laws, by virtue of the realization of certain specified antecedent conditions. [Hempel and Oppenheim 1948, pp. 135–6]

Explanations of this type are relatively uncontroversial. But while the Deductive–Nomological model may do a fine job capturing the key components of some such explanation, this does not necessarily mean that any account satisfying these characteristics should count as a genuine scientific explanation. For among the various problems that the Deductive–Nomological model has been found to suffer, insufficiency is one of the best-known.[7] Consider for example that the height of a flagpole could be explained by including the angle of the sun and the length of the flagpole's shadow as antecedent conditions, and then appealing to laws of geometry or optics. This would subsume the flagpole's height under a (set of) covering law(s), but surely the height of a flagpole cannot be explained by the length of its shadow.[8]

Wesley Salmon recognized that a common feature of many counterexamples to the Deductive–Nomological model was that they exploited its insufficient treatment of causation. In his first attempt to resolve the issue, Salmon promoted what he called the 'Statistical Relevance' model of explanation [1971], which was relatively short-lived.[9] Not long after, Salmon abandoned this model of explanation in favor of the better-known

[7] Other problems are discussed in any of the sources listed in footnotes 4 and 5, above.
[8] Salmon gives credit to Sylvian Bromberger for devising the counterexample, but claims that it was never published [1992, p. 21; 1998, Ch. 19]. In fact Bromberger published a very similar example involving the height of the Empire State Building [1966, p. 92].
[9] A summary and references can be found in the overview by Woodward [2010].

'Causal Mechanical' model [1984].[10] In short, a causal–mechanical explanation of some particular event involves providing a narrative account of the causal processes and causal interactions leading up to the event.

> The basic idea—stated roughly and briefly—is that an intersection of two processes is a *causal interaction* if both processes are modified in the intersection in ways that persist beyond the point of intersection, even in the absence of further intersections. When two billiard balls collide, for instance, the state of motion of each is modified, and those modifications persist beyond the point of collision. A *process* is *causal* if it is capable of transmitting a mark—that is, if it is capable of entering into a causal interaction. [Salmon 1998, p. 71]

When the explanandum is incorporated into a narrative of some such processes and interactions, the event is said to have been fit into a 'causal nexus', and therefore given a causal–mechanical explanation. In doing so, a particular type of understanding is provided, relating primarily to the underlying mechanisms of the explanandum event. As Salmon explains, the aim of causal–mechanical explanations is "to exhibit the ways in which nature operates; it is an effort to lay bare the mechanisms that underlie the phenomena we observe and wish to explain" [1998, p. 71].

During the time leading up to Salmon's introduction of causal–mechanical explanations, the 'Unificationist' model of explanation was promoted by Michael Friedman [1974]. Since then, Philip Kitcher has developed the view into the form by which it is now known [1976; 1981; 1985; 1989]. The details of the Kitcher's Unificationist model would take us beyond the scope of this summary, but the gist of a unificationist explanation is simply that some connection is established between features of the world that were previously thought to be entirely independent.

> The basic idea of the unificationist account is that scientific explanation is a matter of providing a unified account of a range of different phenomena. [...] Successful unification may exhibit connections or relationships between phenomena previously thought to be unrelated and this seems to be something that we expect good explanations to do. [...] Paradigmatic examples include Newton's unification of terrestrial and celestial theories of motion,

[10] The model was edited and updated in Salmon's later works [1994; 1997].

> and Maxwell's unification of electricity and magnetism.
>
> [Woodward 2010, p. 45]

By the examples that Woodward gives, one benefit of unificationist explanations should be clear—namely, that they allow a smaller number of independent assumptions to be made in accounting for natural phenomena.[11] And indeed, specifying those assumptions, and the explanations that they collectively provide, is what Kitcher takes to be the main goal of any model of explanation.

> [S]uccessful explanations earn that title because they belong to a set of explanations, the *explanatory store*, and the fundamental task of a theory of explanation is to specify the conditions on the explanatory store. Intuitively, the explanatory store associated with science at a particular time contains those derivations which collectively provide the best systematization of our beliefs. [Kitcher 1989, p. 429]

Thus unificationist explanations provide a different sort of understanding than causal–mechanical explanations. While the latter "lay bare the mechanisms that underlie the phenomena", the former shows various phenomena to be different manifestations of one and the same physical process, or at least to be in accord with one and the same (set of) laws(s).

The Causal–Mechanical and Unificationist models are two main players in ongoing work on scientific explanation. The difference in the type of explanation that each provides raises the question of whether any one model of scientific explanation could cover all the sciences, or even physics alone. In his review of the then-current status of scientific explanation at the close of the 20th century, Salmon argues that one of the main lessons to be learned is that we should not expect to find one, or even just a few, models of scientific explanation that will have universal application [2011, p. 316]. One of his key reasons for his saying so was first outlined by Kitcher [1985], who explained that all contemporary models of explanation can be categorized as either 'bottom-up'

[11] This is discussed further by Salmon [1998, pp. 69–70], who points out that one of Friedman's original goals in promoting unificationist explanations was to provide an account for laws themselves, which was a stumbling block for the Deductive–Nomological model.

or 'top-down'. Hempel's and Salmon's models fit into the former category, in that they "believe that there can be local explanations of particular facts" and "try to work up from there to more fundamental causal mechanisms or more comprehensive theories" [2011, p. 317]. Friedman's and Kitcher's fall into the latter, according to which "to explain something is to fit it into a global pattern" and "what qualifies as a law or a causal relation is determined by its place in the simplest and most comprehensive theories" [2011, p. 317].

Toward a Broader Framework

From the outline above, it should be clear that Einstein's principle–constructive distinction cannot be taken as a comprehensive framework for the explanatory power of physical theories. For while certain connections can be drawn between causal-mechanical explanations and those provided by Einstein's constructive theories, it would be much harder to argue that Friedman and Kitcher's 'top-down' unificationist models would fit Einstein's criteria for a constructive theory. At the very least, it would seem that principle-theories should be ascribed some explanatory power, even if of a different nature than that of constructive theories. Along those lines, Flores has argued that although what Einstein calls a 'constructive theory' can indeed provide what Kitcher and Salmon call a 'bottom-up' explanation, those explanations involve laws that are themselves given top-down explanations by principle-theories [1999, §4]. And so even if they are not capable of providing explanations for particular phenomena, it should be granted that principle-theories can at least provide explanations for certain laws themselves.[12] Also along similar lines, Dennis Dieks has compared the explanations provided by Einstein's constructive theories and principle-theories with bottom-up and top-town explanations, respectively. He argues against the search for a "uniquely best explanatory scheme", and in favor of acknowledging "a plurality of possible physical explanations and ways of understanding physical processes" [2009].

[12] See also footnote 11, above.

Others, however, would argue against making so strong a connection between Einstein's principle–constructive distinction and Kitcher's top-down and bottom-up explanations. Instead, principle-theories like Einstein's might be understood as playing another sort of explanatory role. Recently, van Camp [2011*b*, §6; 2011*a*, §4] has argued that Flores's analysis of Einstein's framework is incomplete for precisely this reason.

> Once it has been established that the [principle–constructive] distinction is best understood in terms of explanatory function, it must be recognized that it cannot simply be a matter of causal-mechanical vs. Kitcher-style unification. Principle or framework theories can do something more than this. [...] It is tempting to lump principle theories into the top-down class of explanation by unification, the most important competing explanatory theory, since many, if not all, successful principle theories do provide such unification. However, functioning as a framework is not only about unifying argument patterns. It can also be about establishing principles [that are] constitutive of the conceptual framework. [van Camp 2011*b*, p. 27R]

Van Camp's first point is simply that in outlining the 'functional' aspect of Einstein's distinction, Flores ought to have given more attention to the explanatory role of the two theory-types. But his main point is that in doing so, one ought to recognize the function that principle-theories play in establishing the framework within which all other explanations—whether principled or constructive—are to be made [2011, p. 29].[13] The idea is tied closely with DiSalle's understanding (mentioned above) of principle-theories as those which provide a "general framework for inquiry".

Either way, it is clear that some framework broader than Einstein's will have to be adopted in order to recognize the diversity of physical theories and the types of understanding that they provide. And while the aim of this section is not to offer a comprehensive framework for physical theories and their explanatory power, it should be noted that a step in the right direction has recently been taken by Mattias Frisch, who has uncovered a quite similar, and yet undeniably more robust framework of theories in some of Lorentz's previously untranslated work.

[13] Van Camp also makes the same point elsewhere [2011*a*, §4].

> In [1900], Lorentz distinguished theories that begin by postulating "general principles" [p. 335; *cf.* 1904, p. 83] or "general laws" [p. 336] from theories that attempt to account for the phenomena by postulating a "mechanism of the appearances" [p. 336; *cf.* 1904, p. 82]. As examples of general principles he cited the principle of energy conservation and the second law of thermodynamics. Such general principles, Lorentz said, express "generalized experiences" [1900, p. 337]. Examples of theories postulating mechanisms, according to Lorentz, are the molecular kinetic theory of gases, Kelvin's vortex theory, and Hertz's mechanics of concealed motions. [...] By "mechanism" Lorentz appears to have meant any underlying (micro-)process independent of its nature.
>
> [Frisch 2005, p. 666]

The similarities between Lorentz's and Einstein's distinctions are striking, and it would seem that Lorentz shared Einstein's preference for constructive theories with regard to explanation. Principle-theories, he claimed, "lead us to desirable results, but will not show us much during the trip" [1900].[14] And yet, Lorentz did still acknowledge their explanatory value in other ways, making his a more pluralist approach to scientific explanation. What's more, Lorentz's definition of principle-theories is more general than Einstein's, since he refers not only to purely phenomenological theories like thermodynamics, but also the principle of energy conservation. For this reason, Frisch argues that Einstein's principle–constructive distinction should be replaced with a tri-partite one, which would include the following:

> (i) mechanism or constructive theories, such as the kinetic theory of gasses;
> (ii) purely phenomenological principles, such as the second law of thermodynamics;
> (iii) general principles or constraints on all (or at least multiple) levels, such as the principle of energy conservation. [Frisch 2011, p. 179]

Frisch goes on to argue for a more pluralistic, or ecumenical view of explanation than that offered by Einstein. Granting differences in the depth or quality of understanding offered by each type of theory, Frisch maintains that it would be "a mistake to deny that principle or purely phenomenological theories can provide any explanations at all"

[14] Another striking similarity can be seen between Lorentz's words here and Bell's description of the 'Lorentzian pedagogy' (p. 26). But see also footnote 41 on page 25.

[p. 179]. And while it may not be fully comprehensive, it would certainly be a better framework within which to evaluate the various approaches to relativity outlined in the preceding chapter. Indeed, as explained in the following section, Frisch and Lorentz's framework is especially useful in sorting out where to position the dynamical approach to relativity within the principle–constructive spectrum of physical theories.

2.2 Einstein's Legacy: A Critical Review

Frisch is right to advocate a broader, more ecumenical framework for the explanatory power of various physical theories. But of course it has been Einstein's, and not Lorentz's, taxonomy of theories and their explanatory power which has set the stage for recent discussions on explanation in relativity. And as it happens, several of the authors introduced in Chapter 1 seem to implicitly acknowledge Einstein's claims, and appeal to various aspects of Einstein's description as a way of arguing that some particular approach to relativity should or should not count as 'constructive'. In this section, it is shown that Balashov and Janssen, Brown and Pooley, and Nerlich all emphasize different aspects of Einstein's principle–constructive distinction in arguing for their own approaches to special relativity over the others'. And with a few qualifications outlined below, the motivation for appealing to Einstein's 'constructive' label is the explanatory power that supposedly comes with it. However, working within the parameters of Einstein's framework comes with certain disadvantages for each party, which are outlined along the way. The upshot is that working within the confines of Einstein's principle–constructive distinction has led some of the parties to misunderstand or misrepresent each other's views, and that some broader framework would be required in order to appreciate the explanatory power of each approach outlined in the previous chapter.

Janssen on Building Up a Picture

Einstein's description of constructive theories begins by referring to their "attempt to build up a picture of the more complex phenomena out of the materials of a relatively

simple formal scheme." In the context of special relativity, we can safely take the complex phenomena in question to be relativistic effects like length contraction and time dilation. And as shown below, building up a model of these phenomena is the characteristic to which Janssen appeals in defending the 'constructive' nature of his geometrical approach to special relativity, introduced briefly in §1.2.

Janssen does not at first warm to Einstein's claims about the explanatory power of principle-theories and constructive theories. In his 2003 collaboration with Balashov, the authors actually argue that principle-theories are in fact explanatory.

> Understood purely as a theory of principle, special relativity explains the phenomenon if it can be shown that the phenomenon necessarily occurs in any world that is in accordance with the relativity principle and the light postulate. [Balashov and Janssen 2003, p. 331]

Indeed, the authors take issue with an earlier claim by Craig [2001] that principle-theories are explanatorily deficient, and they dedicate one section of their paper to defending the idea that Einstein's 1905 work and even thermodynamics itself provide perfectly adequate explanations of the respective complex phenomena [2003, §5]. But after Brown and Pooley's reply [2006], in which this particular point of Craig's is defended, Janssen concedes the point [2009, fn. 27]. And then, in order to defend the explanatory power of his geometrical approach to special relativity, Janssen must defend its 'constructive' nature.[15] There, it becomes clear which part of Einstein's description Janssen latches onto: he maintains his and Balashov's description of constructive theories as those that "get at the underlying *reality*" of a phenomenon by providing a model that gives an empirically adequate description of it [2003, p. 331].

> Whereas theories of principle are about the *phenomena*, constructive theories aim to get at the underlying *reality*. In a constructive theory one proposes a (set of) model(s) for some part of physical reality (e.g. the kinetic theory modeling a gas as a swarm of tiny billiard balls bouncing around in a box).

[15] It should be noted that in his critique of the dynamical approach, Janssen refers to Einstein's principle–constructive distinction as a red herring [2009, p. 38L]. In its place, he offers the kinematics–dynamics distinction outlined above (§1.2). However, this does not stop him from defending the 'constructive' nature of his geometrical approach.

> One explains the phenomena by showing that the theory provides a model that gives an empirically adequate description of the structural features of reality. Consider the phenomenon of length contraction. [...] A constructive version of the theory [...] explains length contraction if the theory provides an empirically adequate model of relevant features of a world in accordance with the two postulates. Such constructive-theory explanations do tell us how to conceive of the reality behind the phenomenon.
>
> [Balashov and Janssen 2003, p. 331]

In particular, they argued that the "empirically adequate model" was Minkowski spacetime itself [p. 331]. Janssen maintains this position in 2009 by arguing that Einstein was wrong to continue to see special relativity as a principle-theory after Minkowski's 1909 geometrization of it. Instead, Janssen understands Minkowski's work to have "turned special relativity into a constructive theory by providing the concrete model for the reality behind the phenomena covered by the principle theory", and that this model "identified the reality underlying the relevant phenomena as Minkowski space-time" [p. 40R].[16]

> The Lorentz invariance that can be derived from the postulates [...] finds its natural interpretation in terms of the geometry of Minkowski space-time. On my definition of the principle–constructive distinction, this interpretation amounts to the constructive-theory version of special relativity. It says that the space-time component of any acceptable model of a world in accordance with the postulates is Minkowski space-time. [Janssen 2009, p. 39R]

Thus Janssen defends the 'constructive' nature of his geometrical approach. However, it is not clear that what Janssen has provided is anything like what Einstein had in mind when he described models that "attempt to build up a picture of the more complex phenomena". For whereas in Einstein's example, complex phenomena are reduced to the behaviors of more basic material entities being governed by particular dynamical laws, it is not the case that Minkowski spacetime 'builds up' the phenomena of length contraction or time dilation in material objects like rods and clocks. This is especially true given that Janssen claims to be speaking from a relationalist's position, according to which

[16] It should be noted that Janssen does not mean to tie his geometrical explanation to spacetime substantivalism. In fact, he is clear about taking a relationalist position [2009, p. 28L, p. 40R].

Minkowski spacetime is not an entity in the first place.[17] A similar point has recently been argued by Frisch.

> It is unclear what, in the case of special relativity, the corresponding more complex phenomena could be that are built out of Minkowski spacetime as simple formal scheme. In particular, the upshot of a spacetime account of length contraction, for example, seems to be the claim that length contraction need *not* be understood as a complex phenomenon, as Lorentz's account has it, but rather can be understood as simple consequence of the geometry of spacetime. [Frisch 2011, pp. 178–9]

And so it would seem that Janssen has taken liberties with the notion of 'providing a model'. In fact, by Janssen's criteria for a constructive theory, it could even be argued that the 'structural' explanation promoted by Dorato and Felline [2010], mentioned briefly in §1.2, which uses Minkowski spacetime as an abstract model without actually assuming its existence, is also a constructive explanation. But surely this too would miss the mark of what Einstein had in mind.

It should also be noted that Janssen's effort to secure Einstein's 'constructive' label prevents him from doing justice to his own perspective. For one of the most striking features of Janssen's perspective is how it provides a kind of unificationist explanation. Consider, for instance, the fact that Janssen argues that the Lorentz covariance of various dynamical laws can be explained in one fell swoop by Minkowskian geometrical structure. That is, in giving an account for the coincidence of various dynamical laws' Lorentz covariance, Janssen claims to make what he calls a 'common origin inference' to Minkowski spacetime.

> In Lorentz's theory as well as in Brown's proposal, it is a brute fact that these different laws are Lorentz invariant. In the orthodox version of special relativity this cosmic coincidence is traced to a common origin, Minkowski space-time. [Janssen 2009, p. 48R]

As Janssen explains [2009, p. 48R], common origin inferences have been a topic of interest in several of his earlier publications [1995; 2002*a*; 2002*b*], and even his collaboration

[17] See footnote 16, above.

with Balashov [2003]. The appeal of a common origin inference comes in providing a single assumption that can allegedly account for the coincidence of all dynamical laws' Lorentz covariance. And this, of course, gives Janssen's geometrical approach the flavor of a unificationist explanation, as described in §2.1, above. And yet, in emphasizing that the geometrical approach provides a 'model' of relativistic phenomena, Janssen either fails to recognize this, or else avoids making the connection.

In fact, in a draft of his 2009 paper, Janssen does try to clarify the style of explanation provided by his geometrical approach.

> The sense in which Minkowski space-time explains Lorentz invariance is not causal but closer to the sense of explanation captured by the old covering-law or deductive-nomological account of explanation. [Janssen 2008, p. 6]

Now, given that Janssen and Balashov had previously given a similar description for the explanatory power of *principle*-theories [2003, p. 332], it is no surprise that Janssen removes the comment in 2009, where he is defending the *constructive* nature of his geometrical approach. Besides, given the problems with this style of explanation pointed out by Brown and Pooley, [2006, pp. 75–6], Janssen had by that point disavowed the covering law model of explanation [2009, fn. 27]. But in its place, Janssen fails to emphasize the unificationist aspects of his geometrical explanation.[18] Perhaps, had Einstein's framework not been so exclusive as to cause Janssen to avoid construing his approach in any 'principled' way, Janssen could at least have presented his approach for all its worth.

Brown and Pooley on Providing Micro-Physical Detail

A second characteristic of constructive theories comes up in Einstein's analogy to the kinetic theory of gases—namely, the provision of detail of the system's micro-physical structure, or of the dynamical laws governing it. This is the characteristic to which Brown and Pooley appeal arguing that geometrical approaches fail to qualify as constructive theories.

[18] Wesley van Camp has also recently observed that Janssen's common origin inference is effectively a unificationist explanation of Lorentz covariance [2011*a*, p. 1101].

Unlike Janssen, Brown and Pooley appear to accept Einstein's claims about the explanatory power of principle-theories and constructive theories from the beginning. It was mentioned above that they endorse a certain one of Craig's claims regarding the explanatory power of principle-theories, and in fact they seek a constructive rendition of special relativity precisely for the sake of its explanatory power.

> What we wish to consider here is the question of the possibility of a fully constructive rendition of SR, and in particular the possibility of a constructive explanation of the 'kinematical' effects associated with length contraction and time dilation.
>
> The issues surrounding this question have been discussed recently by Balashov and Janssen [2003]. [... They] take explicit issue with two claims that they attribute to Craig [2001]: (i) that SR in its 1905 form fails to provide a theory-of-principle explanation of phenomena such as length contraction and, (ii) that theory-of-principle explanations in general are deficient [2003, p. 332]. We side with Craig on both counts, although it should be stressed that we endorse (i) for reasons quite different to those that motivate Craig.
>
> [Brown and Pooley 2006, pp. 74–5]

As for what makes a theory 'constructive', it is clear that they see this to be the provision of detail in either a system's micro-physical structure or the dynamical laws that govern it.

> What is definitive of this [our] position is the idea that constructive explanation of 'kinematic' phenomena involves investigation of the details of the dynamics of the complex bodies that exemplify the kinematics. [...]
>
> We have been arguing that the truly constructive explanation of length contraction involves solving the dynamics governing the structure of the complex material body that undergoes contraction.
>
> [Brown and Pooley 2006, pp. 81–2]

Indeed, this definition of constructive theories is clear in the authors' endorsement of Bell's work, where it is claimed that a certain depth of understanding comes only from taking the 'longer road' that involves an investigation of the details of a system's dynamical laws. It is also clear from the authors' arguments against the constructive status of Balashov and Janssen's geometrical explanations: "They do not, *directly*, concern the

details of the bodies' microphysical constitution" [2006, p. 79].[19] That said, it should be noted that despite the contrast shown here, Brown and Pooley actually cite Balashov and Janssen's description of constructive theories—as theories that get at the underlying reality—approvingly [2006, fn. 8]. But of course, this is unsurprising when the term 'underlying' is emphasized, given their emphasis on micro-physical detail.

It turns out, however, that Einstein's principle–constructive framework also causes some trouble for Brown and Pooley. Their commitment to the explanatory power of constructive theories, combined with their emphasis that constructive theories provide micro-structural and dynamical detail, actually leaves the dynamical approach outside of the coveted category. And so despite their claims to be in search of a fully constructive approach to special relativity, Brown and Pooley shift gears somewhat suddenly toward the end of their paper:

> [...] one might be tempted to deny that explanations which appeal to an explanans as non-concrete as the *symmetries* of the laws are genuinely constructive explanations. In other words, it turns out that there are even fewer contexts than one might have at first supposed in which length contraction stands in need of a constructive explanation.
>
> [Brown and Pooley 2006, p. 82]

As a result, Frisch understands the dynamical approach to land "somewhere in between an explanatory theory and one that simply fails to explain" [2011, p. 180R].[20] And this is understandable, given that Frisch also latches onto the same defining features of constructive theories as Brown and Pooley—namely, as "constructive micro-theories, that aim to explain by positing detailed microphysical mechanisms" [2011, p. 180L]. In defense of the dynamical approach, Frisch proposes to replace Einstein's framework with a more pluralist one, along the lines of Lorentz's *tri*-partite classification of theories. As mentioned above (p. 46), the latter recognizes the explanatory power of both purely phenomenological principle-theories like Einstein's, as well as the more general 'restricting

[19] The emphasis on micro-physical and dynamical detail is also found in Brown's text [2005*b*, pp. 89–90].
[20] Frisch describes the dynamical approach as a constructive theory from the start [pp. 176–7], though he does of course recognize that Brown and Pooley's dynamical approach does not actually fit into either category of Einstein's framework [p. 179L].

principle' theories like the dynamical approach.

But while Frisch is right to find a better home for the dynamical approach in Lorentz's classification, two points should not be overlooked. First, it is not clear that Brown and Pooley mean to maintain their loyalty to Einstein's framework. The quote above shows that they recognize the dynamical approach to lack what they understand to be the defining features of a constructive theory, and yet they do go on to advocate the dynamical approach all the same. Indeed, Brown makes this clear in the way the dynamical approach is presented in his book.[21]

> The main lesson that emerges, as I see it, is that the special theory of relativity is incomplete without the assumption that the quantum theory of *each* of the fundamental non-gravitational interactions—and not just electrodynamics—is Lorentz-covariant. [...] The *full* version of this pedagogy involves providing a constructive model of the matter making up a rod and/or clock and solving the equations of motion in the model.
>
> [Brown 2005*b*, pp. 4–5; *cf.* pp. vii–viii]

Second, while the authors deny that geometrical explanations are constructive, they do grant that they are nonetheless explanatory in some contexts [2006, pp. 78–9]. This too shows that the authors do not mean to reserve explanatory power strictly to constructive theories. So while it should certainly be acknowledged that their paper is not as clear as it could be on the constructive status of the dynamical approach, Frisch has not necessarily saved them from as awkward a position as it seems.[22]

Finally, the same tension in Brown and Pooley's work is responsible for Janssen focusing his criticism on fully-detailed constructive theories, rather than on the dynamical approach properly understood. That is, because Brown and Pooley promote the search for a fully constructive theory, and only briefly point out how the dynamical approach

[21] Brown makes a similar observation about Einstein's later understanding of special relativity [2005*b*, §8.4.1]. And in another publication of the same year, Brown and Timpson refer to the dynamical approach as a "semi-constructive approach to relativistic kinematics" [2006, p. 37]. Regarding similar comments in Brown's text, Hagar [2008] also makes the point, quoting Brown [2005*b*, p. 147] in calling the dynamical approach a "restrictive *structural* principle" instead of a fully constructive theory: "Brown's big principle is neither a "no-go" phenomenological claim nor is it fully constructive, hence it will not suffice to pin-point the difference between the constructive and the principle approaches to STR" [p. 536].

[22] Note that while Frisch [2011] references incorrect page numbers when quoting Brown and Pooley [2006], the quotations appear to be accurate otherwise.

does not actually fit that criteria, many of Janssen's criticisms of their work are actually focused upon a fully constructive theory, involving the full details of the dynamical laws. The misunderstanding can be traced through Janssen's arguments, which actually support Brown and Pooley's claim that the details of dynamical laws should not play a role in explaining relativistic phenomena. In fact, Brown and Pooley thank Janssen for reminding them of this point [2006, fn. 23]. That many of Janssen's criticisms are misdirected can be seen especially clearly in an analogy that Janssen makes, in which Brown and Pooley are described as explaining the rectangular shape of paintings by appealing to something about the individual paintings themselves, whereas Janssen would explain the shape of paintings by appealing to "factors over and above the individual paintings" like "the practicalities of making a canvas by stretching cloth over a [rectangular] frame" [2009, pp. 27–8]. But a more accurate depiction of the dynamical approach, taking into account Brown's description of it as a "restrictive principle" or a "*super law*" [2005*b*, p. 147], would have Brown and Pooley explain the shape of paintings by some "factors over and above the individual paintings" as well. Indeed, the two parties are largely in agreement about the fact that Lorentz covariance—and not the full details of the dynamical laws—is the proper explanans, and the two parties make a number of strikingly similar comments along these lines.[23] But despite all of this, the authors' attention to different aspects of Einstein's description of constructive theories is clearly what leads to Janssen's misrepresentation of the dynamical approach.

Nerlich on Reducing the Phenomenon

A final characteristic of constructive theories comes up even more more directly in Einstein's analogy to the kinetic theory of gases—namely, reducing the complex phenomenon. Granted, reduction will in some cases be accomplished by providing detail of the system's micro-physical structure, as in Brown and Pooley's definition of a constructive theory. However, in defending the 'constructive' nature of his geometrical approach, Graham

[23] Compare [Janssen 2009, p. 39L] with [Brown and Pooley 2006, p. 82], and [Janssen 2009, pp. 40R–41L] with [Brown 2005*b*, p. 90, p. 147]. Frisch has made a similar observation [2011, pp. 180–1].

Nerlich [2013] argues that both of the above parties have misunderstood this aspect of Einstein's distinction.

Unlike Janssen, Nerlich does not put any emphasis on Minkowski spacetime's provision of a 'model' for relativistic phenomena; unlike Brown and Pooley, Nerlich does not define constructive theories by their provision of micro-physical or dynamical detail. Instead, he emphasizes the kinetic theory's reduction of temperature and pressure.

> [The] core criterion of a constructive explanation is that it is reductive: phenomenal things and properties are displaced at a new level by more basic theoretic ones. That the new entities and properties may, but need not, be micro physical is a merely accompanying feature and the reducing theory may, but need not, be statistical. [...] Constructive explanations are not causal.
>
> [Nerlich 2013, p. 77]

And while it was granted above that Brown and Pooley's attention to micro-physical detail entails some kind of reduction, Nerlich puts particular emphasis on reducing the *ideology* of a theory.

> [A] constructive explanation for the contraction of a moving rod [...] would be a theory in which concepts of length, motion, rod, contraction and clock-rate are replaced by other entities and properties that explain how and why the contraction of rods in motion is phenomenal. There is no such *microphysical* theory consistent with Special Relativity. It is hard to imagine one that would be.
>
> [Nerlich 2013, p. 78]

With constructive theories thus construed, Nerlich feels that neither of the approaches above meet Einstein's criteria. A large part of his criticism of Brown and Pooley's work is the authors' endorsement of Bell's work, whose (allegedly) fully detailed dynamical approach does not reduce the ideology of the phenomenon, and whom Nerlich takes to be offering a causal explanation, which he distinguishes from a constructive explanation [pp. 81–83]. His main criticism of Balashov and Janssen's work is the authors' emphasis on models instead of reduction [fn. 3, p. 78]. On the other hand, his positive case for the constructive nature of his own geometrical approach to special relativity is the claim that Minkowskian geometry provides a new ontology *and* ideology of the relativistic

effects in terms of a four-dimensional picture [p. 90].[24] In particular, when relativistic effects are recognized as artifacts of frame-dependent, three-dimensional projections of the world's actual four-dimensional ontology and ideology, Nerlich sees relativistic effects to be given a reductive and non-causal explanation. And this would mean that his geometrical explanation meets the criteria he takes Einstein to have suggested in his example of the kinetic theory of gasses.

Now, insofar as Einstein's search for an analogue to the kinetic theory of gasses is worth completing, it might be argued that Nerlich has provided the first defense of a geometrical approach with the right aim. For whereas Balashov and Janssen take advantage of the ambiguity in the criterion of 'providing a model', Nerlich focuses on specific aspects of Einstein's illustration involving the kinetic theory of gasses. As in the quotations above, Nerlich takes the example to show that ontological and ideological reduction are key elements at play, and that causality or micro-physicality are common—but not necessary—features of a constructive explanation. However, Nerlich's agreement with Einstein's constructive-principle framework leads him to defend his geometrical approach in ways that arguably don't emphasize its strengths. For instance, Nerlich's geometrical explanation is like Janssen's in providing a unificationist style of explanation that appears in Friedman's defense of space-time realism [1983].[25] But in an attempt to describe it as a constructive theory, Nerlich emphasizes instead the allegedly reductive aspects of dimensional projection.

> There are 4-dimensional objects, worldlines and worldsolids set in a space-time of appropriate geometrical structure. This re-conceives the motion and contraction of moving rods as phenomena resting on the relative orientations of 4-dimensional entities and their 4-dimensional properties as these occur in Minkowski spacetime. [...] These new 4-dimensional objects do not move, they do not contract with motion therefore, and the slowing of clocks is a matter of the proper time metric along their worldlines. The new things and new properties yield a reductive explanation of the phenomenal properties of phenomenal rods and so on, as these appear in frames. [Nerlich 2013, p. 91]

[24] Nerlich gives a similar introduction to relativistic ontology and ideology elsewhere [2010].
[25] See Nerlich's comments along these lines [2013, pp. 90–1].

Yet insofar as dimensional projection counts as a form of reduction, it is not clear that it achieves the kind of reduction that Einstein illustrated with the kinetic theory of gasses.

Second, in order to present his geometrical explanation as 'constructive', Nerlich must argue against causality and micro-physicality as necessary characteristics of a constructive theory. This comes up briefly in the quotations above. To show that constructive theories are not necessarily micro-physical, Nerlich gives an example of a micro-physical explanation for the color of blood which, while micro-physical, is not reductive [p. 77]. His point is apparently to show that micro-physical explanations are not necessarily reductive, but of course this says nothing about whether a constructive-theory explanation is a micro-physical one. Indeed, the two examples that Nerlich offers—one being Einstein's, and another borrowed from Arthur Eddington—are both micro-physical [p. 77], which would require that Nerlich give some argument as to why we should take the reductive aspects of those explanations instead of the micro-physical aspects as key. The same would go for Nerlich's claim that constructive-theory explanations need not be statistical [p. 77]. Without clearer argument for honing in on the reductive aspects of these explanations, these two claims amount to mere assertions. But these assertions would not be necessary if Nerlich were not defending the 'constructive' nature of his geometrical approach.

2.3 Constructive Relativity and Its Critics

So far, this chapter has introduced Einstein's principle–constructive framework, showed its deficiencies as a comprehensive framework of physical theories, and outlined some related problems in the recent literature on explanation in relativity. But now, leaving Einstein's principle–constructive distinction behind, the rest of this section focuses primarily upon recent criticisms of the dynamical approach that focus upon more metaphysical issues.

It was mentioned above (p. 13) that part of Einstein's concern with his 1905 introduction to special relativity was the way that rods and clocks were taken as primitive entities. Avoiding this 'sin', as he called it [1949*a*, p. 61], is actually an important part of the dynamical approach to special relativity. For it is in conceiving of rods and clocks

as complex material objects, comprising more fundamental entities governed by certain dynamical laws, that Brown and Pooley latch onto Lorentz covariance in accounting for rods' length contraction and clocks' time dilation.[26] And because the authors therefore take Lorentz covariance to account for the fact that the best codification of rods' and clocks' behavior is Minkowski geometry, it should be unsurprising that some authors have understood Brown and Pooley to defend the idea that Lorentz covariance accounts for Minkowski geometrical structure without the need to stipulate any other geometrical structure or geometrical concepts. Two examples are given below, followed by a clarification as to how the dynamical approach ought to be properly understood.

Why Constructive Relativity Fails

An unequivocal account of explanation is elusive enough that some have chosen to avoid the topic altogether. Norton [2008] has recently taken this route by criticizing the dynamical approach on metaphysical grounds, rather than explanatory grounds. And while Frisch has rightly pointed out [2011, fn. 4] that some of Norton's criticisms actually hinge upon explanatory issues, those are not the focus here.[27]

In short, Norton's concern with the dynamical approach is based on the assumption suggested above: that the dynamical approach seeks to 'construct' Minkowski geometry without (m)any other assumption(s) of primitive geometrical concepts. That is, Norton describes the 'constructive' approach to relativity as being committed to the following claims.

> that specially relativistic spacetime geometry "can be inferred from the properties of matter without recourse to spatiotemporal presumptions or with few of them"
>
> that spacetime theories "are essentially matter theories [and] familiar spacetime geometries are dependent on properties of matter and induced by them"

[26] To be clear, Brown understands rods and clocks to be atomic configurations governed by Lorentz-covariant dynamical laws [2005*b*, p. 4]. That is, he takes them to be *solutions* of the dynamical equations, and he takes those equations' Lorentz covariance to be that which accounts for rods' and clocks' behavior being encoded by Minkowski geometry. See also Brown's later comments along these lines [2009, §3.2].

[27] Norton ends up defending Janssen's common origin inference argument [2008, §6].

> that "spacetime results from matter in the sense that spacetime structures are all really properties of matter; or, more carefully, spacetime supervenes on matter" [Norton 2008, pp. 821–2]

But according to Norton, the constructivist project would require that one "antecedently presume the essential commitments of a realist conception of spacetime", which can only be avoided "by adopting an extreme form of operationalism" [p. 821]. These two arguments are outlined in turn. This section offers only a partial response, as the following chapters are largely motivated by Norton's concerns.

As for antecedently presuming the essential commitments of a realist conception spacetime, Norton's understanding of a 'realist conception' should be clarified. In short, Norton has in mind what might also be called a substantivalist's geometrical approach to special relativity, which involves: i) the existence of a four-dimensional, non-Euclidean spacetime; ii) the point-like parts of that spacetime being related by spatiotemporal intervals; and iii) the dynamical laws that govern material objects being *adapted* to the independent geometry of that spacetime [p. 823]. And to demonstrate how the constructivist project must accept some or all of these, Norton outlines the steps in a hypothetical constructivist project [§§2–3] and claims to show that it requires the constructivist to presume a notion of spacetime coincidence.

> One cannot recover the unique four-dimensional spacetime [...] of the realist's conception of spacetime except by presuming it at the start. In effect, [the project] presumed a single spacetime with topological properties and standard coordinates (x, y, z, t). This does not augur well for constructivists who seriously believe that Minkowski spacetime is a 'glorious non-entity.' [Norton 2008, pp. 828–9]

Norton takes this assumption to be a form of the first realist conception of spacetime—that Minkowski spacetime exists independently. Then, if Norton's point is granted, he can make two further arguments. First [§5], that if the standard coordinates described above are in fact coordinates of an independently existing spacetime, then Lorentz covariance is "not simply the expression of a property of matter", but rather "the expression of

a property of matter and spacetime taken together" [p. 829]. And this, Norton claims, amounts to a form of the third component of the realist conception—that the dynamical laws are adapted to the geometry of spacetime (i.e. that the arrow of explanation runs from geometry to dynamics). Second [§6], Norton argues that if his first argument holds, then spatial distances and times elapsed are properties of spacetime, not matter. Furthermore, he claims that in order to avoid this conclusion, the constructivist must take the 'extremely operationalist view' in arguing that for empty or homogeneous areas of spacetime, there are no distances or times elapsed, as none are actually measured.

> If times elapsed are to supervene upon matter, or more vaguely to be a result of the properties of matter, then the absence of any change in the matter entails that there is no change in times elapsed. [Norton 2008, pp. 831–2]

Now, it should be clear from the presentation above that Norton's second and third arguments follow only if the first is granted. The first is discussed further below, but in the meantime the second and third can be addressed briefly. Regarding the second argument, it should be pointed out that arrow of explanation from Minkowski spacetime to Lorentz covariance is only supported if Minkowski spacetime is taken as primitive. But Norton's first argument has only shown that in order to make use of a coincidence relation, some spacetime with topological properties must be assumed. The rich structure of Minkowski spacetime far surpasses mere topological structure, and so it is not clear how we pass from topology to Minkowski geometry. What's more, it is not clear that what needs to be assumed is some topologically structured *entity* called spacetime. For if the material objects in question were granted some primitive topological structure, then the same coincidence relation could be assigned to the fundamental entities without calling into existence substantival spacetime.[28] How well this kind of move would fit in the dynamical approach is addressed below.

Regarding the third argument, two points made recently by Oliver Pooley are worth repeating here. First, with regard to empty regions of spacetime, there is no problem

[28] Incidentally, it is be argued in Appendix B that some such coincidence relation can be reduced to topological indiscernibility.

to be faced—"for the constructivist there is *literally nothing* in an empty region and so nothing whose geometrical properties might be indeterminate" [2013, p. 573]. As for homogenous regions of matter, Pooley reminds us that while certain geometrical facts will be indeterminate, this is at most a peculiarity of the constructivist's position, and not anything devastating to the constructivist project—i.e. it would not be so hard a bullet for the constructivist to bite.

Again, both Norton's second and third arguments hinge upon his first. And while Norton's first argument goes too far in insisting that a substantival spacetime be assumed, it does seem right to say that the constructivist would need to presuppose at least some kind of topological structure on material objects. That Norton takes this to be devastating to the constructivist project only shows that he understands the goal to have been to recover Minkowski geometrical structure without presupposing any geometrical structure whatsoever. But, that the dynamical approach should not be understood in this way is addressed further at the end of this chapter.

The Primacy of Geometry

Amit Hagar and Meir Hemmo [2013] have also raised a challenge for 'dynamical approaches' (generally speaking), which focuses on more metaphysical issues. But whereas Norton focuses upon aspects of the realist's conception of spacetime, Hagar and Hemmo focus specifically on the ideological component of the project. That is, rather than arguing that the dynamical approach must assume some independently existing spacetime, the authors argue that "in current spacetime physics there can be no dynamical derivation of *primitive* geometrical notions such as length" [p. 357].

To begin, the authors outline their understanding of a dynamical approach to relativity.

> *Dynamical* approaches to relativity [... view] the dynamics as primitive and as preceding geometry, and attempt to derive the Minkowski spacetime structure from the dynamical behavior of rods and clocks without *any* further input of primitive geometrical concepts. [Hemmo and Hagar 2013, p. 357]

And to be clear, they have particular geometrical notions in mind.

> [The] issue at stake is whether one can dispense altogether with *primitive* geometrical notions such as length, area, volume, etc., in the dynamical attempts to derive the geometrical structure of spacetime. That is, the issue at stake is whether one can derive the notion of an interval from the dynamics alone. [Hemmo and Hagar 2013, p. 358]

But the problem, as they see it, comes up in Einstein's correspondence with Swann, in which Einstein argues that for any theory that does not treat rods and clocks as independent physical objects, some notion of length must be included in the theory's primitive ideology [pp. 358–9]. Thus any dynamical approach, by their definition of the term, must do precisely what it aims not to do.

In support of the claim, Hagar and Hemmo focus primarily upon several examples: Swann's approach to special relativity [§4] and loop quantum gravity and the causal set approach [§5]. Yet unlike in Norton's paper, no attention is given to Brown and Pooley's collaborations, or to what Norton calls the "abbreviated constructive project" outlined in Brown's text [2005*b*, Ch. 2]. Instead, they commit a footnote to acknowledging that Brown and Pooley do not defend the rather 'strong' dynamical project outlined above, but should rather be understood as defending a somewhat weaker claim. According to the weaker claim, "the symmetries of the geometry of spacetime are to be derived from features of the dynamics" [2013, fn. 3], but primitive notions like length or area can indeed be taken as primitive.[29] And so to some extent, the authors' concern over taking a notion of length as primitive is not so much a challenge to *the* dynamical approach as it is to other 'dynamical' approaches like those just mentioned. In that case, one is also left to wonder for whom the arguments are relevant. For if not Brown and Pooley, who does in fact advocate the 'strong' dynamical approach against which the authors argue? It would seem that their target is only Swann himself, at least as they interpret his work.

What's more, it is not clear that the authors have discharged Brown and Pooley for the right reasons. Above, their reasoning seems to be that the dynamical approach should be allowed to take geometrical concepts as primitive. But it would surely be wrong to say

[29] This 'weaker' dynamical project is contrasted with the stronger in two places [2013, p. 358L, p. 359R].

that the 'constructivist' could take, say, spatiotemporal intervals as primitive. Taking Minkowski interval relations as primitive, for instance, would amount to a 'flat-footed' relationalist approach to special relativity that has already been explored at length and found to have its limitations.[30] But more importantly, from such a primitive relation, Minkowski spacetime geometry would not be left to be accounted for by the symmetries of the dynamical laws. And so in fact the dynamical approach does attempt to do what Hagar and Hemmo take Einstein to say can't be done: to account for Minkowski spacetime structure without treating rods and clocks as independent entities, and also without assuming any primitive notion of length or interval. Rather, what separates the dynamical approach from the 'strong' dynamical project is that so long as Minkowski geometrical structure is accounted for by the symmetries of the dynamical laws, some other sub-metrical geometrical structures could in fact be assumed. This is discussed further below.

New Work for the Dynamical Approach

As discussed above, Norton is right in pointing out that in addition to the explanatory claims outlined above (§1.3), the dynamical approach would appear to involve an implicit ontological claim, which is evident in the title of Brown's second collaboration with Pooley: *Minkowski Space-Time, a Glorious Non-Entity*. That ontological claim is also acknowledged by Huggett [2009], by Nerlich [2013, Ch. 5], and in more recent, independent work by Pooley [2013, §6.3.2]. And so while the primary focus of Brown's work is not the ontological status of spacetime, the dynamical approach to relativity is arguably best construed as a form of relationalism, according to which spacetime does not exist primitively. For otherwise, it could not be said that the symmetries of the dynamical laws somehow account for its geometrical structure.[31] But more importantly, it should also be

[30] See the references in footnote 35 on page 65.

[31] This assumes that any substantival spacetime would have a primitive geometrical structure. But of course it should be noted that, even if kinematical effects and inertial motion are explained by an appeal to the symmetries of the dynamical laws, the dynamical approach might be compatible with a form of substantivalism according to which spacetime has less geometrical structure than that which best codifies the behavior of material objects. See also the discussion surrounding footnote 55 on page 33, and the

pointed out that this explanatory claim does indeed come with an implicit *ideological* claim: that material objects' primitive properties do not constitute a Minkowskian geometrical structure—for instance: that material objects do not stand, primitively, in the same chronometric relations that they would if existing in substantival Minkowski spacetime.[32] Otherwise, it could not be said that the symmetries of the dynamical laws somehow account for the material world's Minkowskian geometrical structure: the latter would be constituted by the fundamental entities' primitive chronometric relations; the fundamental entities would instantiate Minkowski geometry on their own.[33] This would not do for the dynamical approach, which describes that geometrical structure as "parasitic on the relativistic properties of the dynamical matter fields" and as "no more than a codification of the behavior of rods and clocks" [p. 9, p. 100].[34] Thus of the three explanatory claims mentioned in Chapter 1, the idea that dynamics explains geometry is what first associates the dynamical approach with relationalism, and is also what illustrates the difference between the dynamical approach and a more 'flat-footed' relationalist approach: in the latter, the material world's spatiotemporal geometry is constituted by the primitive properties (chronometric relations) of material objects; in the former, it supervenes upon the symmetries of the dynamical laws.[35]

That said, the challenges outlined above go too far in assuming that *all* geometrical structures or geometrical concepts must supervene upon the symmetries of the dynamical laws. As Pooley has recently reminded us, "the [dynamical approach] project was to reduce chronogeometric facts to symmetries, not to recover the entire spatiotemporal nature of the world from no spatiotemporal assumptions whatsoever" [2013, p. 573]. Still,

longer discussion in §6.1.

[32] As explained in footnote 2 on page 1, the terms 'ontology' and 'ideology' are used as by Quine. (See also the quotation on page 66, below.) Quine notes that "[as] a subdivision of Ideology there is the question of what ideas are fundamental or primitive for a theory, and what ones are derivative" [1951, p. 14]. This distinction is critical throughout the following discussion.

[33] This point is also made by Huggett [2009, p. 417]. It goes against a suggestion by Frisch that, so long as Minkowski geometrical structure is assigned to the material bodies, and not to an independently existing spacetime, Norton's challenge can be avoided [2011, fn. 1].

[34] See also Brown's Chapter 8 [2005*b*, pp. 142–3], or his collaboration with Pooley [Brown and Pooley 2006].

[35] The viability of such a 'flat-footed' relationalist approach to Minkowski spacetime structure is explored by Earman [1989, §6.10], Maudlin [1993, §5], and Pooley [2013, §6.1; 2014, §3.1.4, Ch. 5].

the challenge is left for the advocate of the dynamical approach: how is it that Lorentz covariance can account for Minkowski spacetime without taking Minkowski spacetime structure as primitive? What geometrical structure *should* be taken as primitive, or what, if any, primitive spatiotemporal relations on material objects?

Insofar as the explanatory issues can be decoupled from the ontological and ideological issues, the following sections focus primarily upon the latter by introducing and considering various ways to 'flesh out' the dynamical approach as a complete metaphysical package of a specially relativistic, material world *sans* substantival spacetime. This kind of project will be familiar to metaphysically-minded philosophers of physics, and has been described well in recent work on related issues by Cian Dorr.

> What does it mean to propound a hypothesis 'about the fundamental structure of the world'? According to a standard approach, stating such a hypothesis involves (i) saying something about the *fundamental ontology*: the entities such that all facts ultimately boil down to facts about them; (ii) presenting a *fundamental ideology*: a catalogue of properties of, and relations among, the fundamental entities; and (iii) stating some *laws* which capture important general patterns in the holding of the fundamental properties and relations.
>
> [Dorr 2011, p. 139]

Thus the question at hand is this: if Minkowskian geometrical structure is to supervene upon the behavior of material objects that act in accord with Lorentz-covariant dynamical laws, then what could be the fundamental entities of such a world, and what could be their primitive properties, such that the material world's spatiotemporal geometry is indeed Minkowskian, and such that this geometry is not constituted by the entities' primitive properties, but rather is accounted for by the symmetries of the dynamical laws? From the discussion above, it should be clear that Minkowski spacetime cannot feature in the list of fundamental entities (i.e. in the 'ontology'), and that chronometric 'interval' relations cannot feature in the list of material objects' primitive properties (i.e. in the 'ideology').[36] And so the dynamical approach is to be an ontologically *and* ideologically

[36] Along similar lines, Brent Mundy has provided a 'physical' axiomatization of Minkowski spacetime involving various three-, four-, six-, and eight-place relations among material bodies [1986]. But because these constitute Minkowski geometry, these relations would be of no use to the constructivist project.

relationalist approach to Minkowski geometry: it must provide a material ontology that excludes Minkowski spacetime, and it must provide a spatiotemporally scant ideology, the spatiotemporal components of which do not constitute Minkowski geometry. Then, with respect to some such metaphysics, it must show Minkowski geometry to be derivative of facts about the symmetries of the dynamical laws. And of course, the latter must somehow be expressed with without tacitly assuming the former.

3 Choosing Coordinates:
Geometry and the Simplicity of Laws

Finding a metaphysics to undergird the dynamical approach is to go beyond Brown's own work on the subject. Already, it has been argued that the dynamical approach is best understood as a form of ontological and ideological relationalism, with the ideological component being implied by the explanatory priority of dynamics over geometry. In this chapter, it is shown that the ideological stance of the dynamical approach gives it a unique position in the empiricist–conventionalist debate on spacetime structure. To begin, Einstein's well-known distinction between 'pure' and 'practical' geometry is introduced in §3.1, and it is argued that the dynamical approach is best understood as a form of 'practical geometry', according to which spacetime geometry is determined by the simplest dynamical laws being written in terms of it. In §3.2, similar ideas are revealed in Minkowski's presentation of the geometrization of special relativity, and also in early work on relativity by Moritz Schlick. Finally, this understanding of the dynamical approach is also shown to be compatible with a related discussion on "the physics of coordinate transformations" in Brown's text [2005*b*, Ch. 2].

3.1 The Dynamical Approach as a Form of 'Practical Geometry'

Recall the main thrust of the dynamical approach to special relativity: that relativistic effects, inertial motion, and even Minkowski geometrical structure are to be explained by a structural property of the quantum theory of matter—namely, Lorentz covariance. As mentioned above (p. 22), the first two claims are the least controversial; many advocates of the geometrical approach would grant that the laws' Lorentz-covariance is explanatory of relativistic effects and inertial motion. However, they would then go on to say that the laws' Lorentz-covariance is explained by the geometry of Minkowski spacetime.[1] Hence the emphasis throughout these sections is on the 'arrow of explanation' between the symmetries of the dynamical laws and those of spacetime geometry.

[1] E.g. Janssen [2009], whose work is outlined above (§1.2, §2.1, §2.2). Maudlin [2012] is another example.

The claim that the arrow of explanation runs from dynamics to geometry is prima facie suggestive of the idea that spacetime structure is an empirical issue—i.e. that it is defined by the behaviour of material objects like rods and clocks. And so while Brown doesn't address the issue directly, there appear to be some potential ties between the dynamical approach to relativity and the epistemology of geometry promoted by logical positivists like Hans Reichenbach and Moritz Schlick. But at the same time, it would seem that advocates of the geometrical approach could also accept the idea that spacetime geometry is an empirical issue. For this would not stop Minkowski geometry from playing the alleged explanatory roles outlined in §1.2. So what does the arrow of explanation from dynamics to geometry have to do with all this? The purpose of this section is to refine our understanding of the dynamical approach by finding its place in the empiricist–conventionalist spectrum of approaches to spacetime geometry.

On Einstein's 'Practical Geometry' and Poincaré's Conventionalism

The idea that spacetime geometry is an empirical issue, and that it should therefore be described as a branch of physics, was popularized by Einstein in a lecture before the Prussian Academy of Sciences, which was published under the title *Geometry and Experience* in the same year [1921]. In the opening paragraphs, Einstein makes an important distinction between what he calls 'pure' or 'axiomatic' geometry and 'practical' geometry.[2] The former is Einstein's name for the formal, axiomatic mathematics itself, which consists of primitive terms that are defined implicitly by the axioms they satisfy. That is, while primitive terms like 'straight line' may conjure up certain spatial concepts, the terms themselves are not inherently spatial; they have no more meaning than that given to them by their satisfaction of the given axioms.[3]

> [Pure] geometry treats of objects which are denoted by the words straight line, point, etc. No knowledge or intuition of these objects is assumed, but

[2] Einstein makes the same distinction elsewhere [1925]. Carl Hempel also gives a very helpful introduction [1945]. Some authors call the former 'mathematical' geometry, and the latter 'physical' or 'applied' or 'empirical' geometry.

[3] As mentioned in footnote 16 on page 83, this understanding of geometry comes from David Hilbert.

> only the validity of the axioms, [...] which are to be taken in a purely formal sense, i.e. as void of all content of intuition or experience. These axioms are free creations of the human mind. All other propositions of geometry are logical inferences from the axioms (which are to be taken in the nominalistic sense only). [Einstein 1921, p. 234]

On this view, pure geometry can make no empirical claims. Instead, its value lies in the certainty of its mathematical claims. From the axioms can be deduced sure consequences; the mathematical propositions of pure geometry are logically implied by whatever axioms are being taken as premises.

Practical geometry, on the other hand, is what results from making some coordination between the primitive terms of pure geometry and actual, material objects. With some such coordination in place, the logical deductions of pure geometry can be translated into hypotheses about the behavior of material objects. Then, when certain axioms of a pure geometry enable a practical geometry that makes accurate empirical claims, it can be said that the geometrical structure represented by the given axioms is true of the material world. In this way, spacetime geometry becomes a branch of physics.

> It is clear that the system of concepts of axiomatic geometry alone cannot make any assertions as to the behavior of real objects of this kind, which we will call practically-rigid bodies. To be able to make such assertions, geometry must be stripped of its merely logical-formal character by the coordination of real objects of experience with the empty conceptual schema of axiomatic geometry. To accomplish this, we need only add the proposition: solid bodies are related, with respect to their possible dispositions, as are bodies in Euclidean geometry of three dimensions. Then the propositions of Euclid contain affirmations as to the behavior of practically-rigid bodies. [...] The question whether the practical geometry of the universe is Euclidean or not has a clear meaning, and its answer can only be furnished by experience. [Einstein 1921, pp. 234–5]

As a branch of physics, practical geometry makes no certain, mathematical claims. Hence Einstein's well-known reflection that "as far as the propositions of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality" [p. 233].

Now, Einstein's purposes in making the distinction are clear: it comes as part of his response to a certain epistemological question about geometry.

> How can it be that mathematics, being after all a product of human thought which is independent of experience, is so admirably appropriate to the objects of reality? Is human reason, then, without experience, merely by making thought, able to fathom the properties of real things?
>
> [Einstein 1921, p. 233]

This is the question of whether geometry can provide us with synthetic *a priori* knowledge. At the time of Einstein's writing [1921], advocates of the neo-Kantian tradition held that it could, and Einstein's response has typically been taken as illustrating his ties with logical positivism. His distinction between pure and practical geometry paints a picture according to which until some coordinations are made between the pure geometrical primitives and the material world, and until some experiments are carried out to test the hypotheses that follow from that coordinatization, 'pure' geometry itself provides no information about the material world.[4] But of course, Einstein's views on geometry were relevant not only to discussions on epistemology; according to Einstein himself, they also played a critical role in his development of general relativity. As he reports, it was only by adopting the perspective of 'practical geometry' that Einstein was able to abandon Euclidean spacetime geometries in the first place.

> I attach special importance to the view of geometry which I have just set forth, because without it I should have been unable to formulate the theory of relativity. Without it the following reflection would have been impossible: in a system of reference rotating relatively to an inertial system, the laws of disposition of rigid bodies do not correspond to the rules of Euclidean geometry on account of the Lorentz contraction; thus if we admit non-inertial systems on an equal footing, we must abandon Euclidean geometry. Without the above interpretation the decisive step in the transition to generally covariant equations would certainly not have been taken. [Einstein 1921, p. 235]

That Einstein put the concept of practical geometry to use in developing general relativity is well-known, especially given the importance that Einstein puts on the thought-

[4]For further discussion, see related work by Friedman [2002], Giovanelli [2012], or Howard [2014].

experiment above.[5] But the fact that Einstein found the perspective so useful does not mean that he went so far as to accept the corresponding epistemology of geometry. In fact, Einstein seems to have favored another perspective.

Later in the same article, Einstein introduces a rival, conventionalist view of geometry, which is best known from its promotion in the work of Henri Poincaré [1905]. In short, Poincaré rejects the coordination of pure geometry's primitive terms with material objects, and his reasons are clear: Einstein's coordination involves rigid bodies, of which there are none. Any potential candidates—Einstein refers to "measuring lines, measuring wands, etc." as examples [1921, p. 234]—are subject to conditions like temperature and pressure. What's more, Poincaré argues that all talk of rigidity must be made in relation to some geometry anyway, meaning that a geometrical structure has to be assumed in any discussion of rigidity. As Einstein explains [p. 236], Poincaré's denial of the coordination makes 'practical' geometry impossible.[6] Then, Einstein seems to agree with Poincaré in principle, although he does go on to suggest that assuming the existence of rigid bodies has its advantages.

> *Sub specie aeterni* Poincaré, in my opinion, is right. The idea of the measuring-rod and the idea of the clock coordinated with it in the theory of relativity do not find their exact correspondence in the real world. It is also clear that the solid body and the clock do not in the conceptual edifice of physics play the part of irreducible elements, but that of composite structures, which must not play any independent part in theoretical physics. But it is my conviction that in the present stage of development of theoretical physics these concepts must still be employed as independent concepts; for we are still far from possessing such certain knowledge of the theoretical principles of atomic structure as to be able to construct solid bodies and clocks theoretically from elementary concepts. [Einstein 1921, pp. 236–7; *cf.* 1949*b*, pp. 676–9]

Einstein made similar claims when writing on the same topic in the years immediately following [1924; 1925]. Years later, Einstein also staged a fictitious debate between

[5] See Friedman's references [2002, pp. 207ff], most notably Stachel's related work [1989*b*].
[6] At least, it prohibits Einstein's particular coordination. Some other coordination not involving rigid bodies could be suggested, instead.

Reichenbach and Poincaré, attributing something like the 'practical geometry' perspective to Reichenbach, and the same opposing view to Poincaré [Einstein 1949*b*, p. 679].[7]

Of course, Poincaré is known for defending the view that if geometry is not determined empirically, it must be chosen as a convention, and that the choice will always be for the simplest geometry, which he takes to be Euclidean geometry [1898, pp. 41–3; 1905, p. 59].[8] But in fact another form of conventionalism could be considered, according to which the choice goes to the simplest overall combination of geometry and physics. This issue comes up on another occasion, in which Einstein responds to Alfred Elsbach's views on space and time, which bear some similarities to those that Einstein attributes to Poincaré, above.[9]

> The stance on these theses depends on whether one grants reality to practical rigid bodies. If yes, then the concept of distance corresponds to something experienceable. Geometry then contains statements about possible experiments; it is a physical science directly subject to experimental proof (standpoint A). If no reality is conceded to practical rigid measurable bodies, the geometry alone holds no statements about experiences (experiments), rather just geometry together with the physical sciences (standpoint B). Physics has hitherto always used the simpler standpoint A and largely owes to it its productivity; it uses it in all its measurements. [...] If, however, one assumes standpoint B, which at the current state of science should be regarded as exaggeratedly cautious, then geometry on its own is not experimentally testable. Then geometrical measurements don't exist at all. [...]
>
> Seen from standpoint B, the selection of geometric concepts and relations is determined only for reasons of simplicity and practicality. [...] But then nothing can be discerned empirically about the dimensions of the space [...] because with this choice of standpoint geometry is not a *complete* physical system of concepts but only part of one. [Einstein 1924, p. 202]

It is in 'standpoint B' that we see an adaptation of Poincaré's conventionalism, such that

[7] Indeed, the approach that Einstein outlined in 1921 has much in common with that of Reichenbach [1928; 1949], whose use of 'coordinative definitions' establishes a link between 'axiomatic' and 'empirical' geometry. Closer comparisons (and contrasts) are offered by Giovanelli [2012] and Howard [2014].

[8] Poincaré's thoughts on the simplicity of Euclidean geometry have not gone unchallenged. Belot gives a brief discussion [2011, pp. 64–6].

[9] Another translation of the following quotation is given by Howard [2014, p. 20]. Einstein also makes the same point elsewhere [1925, pp. 234–6].

the criterion of simplicity is transferred from the choice of geometry to that of geometry and dynamical laws together. Indeed, this approach is commonly cited as a third way out of the Einstein–Poincaré controversy: it is suggested by Carl Hempel [1945, p. 15], by Einstein's imaginary Reichenbach [Einstein 1949*b*, p. 678], by Michael Friedman [2002], and more recently by Thomas Oberdan [2013].

Configuring the Dynamical Approach

It can now be asked where the dynamical approach fits in among the perspectives outlined above. To begin, it should be pointed out that if emphasis is placed on the explanatory claims of the dynamical approach, then one might be led to think that Brown is advocating something like Einstein's practical geometry. Along these lines, Bradford Skow makes the following observation.

> Obviously, the symmetries of the laws do not cause spacetime to have a certain geometrical structure. So what kind of explanation is this? Maybe Brown thinks that the symmetries of the laws explain why spacetime has the structure it does because the spacetime metric is somehow analyzed in terms of the laws. In several places his discussion suggests that, on his view, a field only gets to be the *metric* field when the dynamics prescribe that material bodies behave in the right sort of way with respect to it (see, for example, pages 100 and 160). Compare this to a Reichenbachian view according to which distance between points of space is analyzed in terms of the behavior of measuring rods—though Brown rejects this view (23), his view does seem to resemble it: both of them make spacetime geometry an extrinsic matter. (Brown's apparent denial that spacetime geometry is "self-standing" and "autonomous" looks like a denial that it is intrinsic (143).) [Skow 2006, p. 4]

Skow's comparison is certainly understandable; the claim that spacetime geometry is to be explained or accounted for by the behavior of rods and clocks would seem to entail its being *defined* by the behavior of rods and clocks. And yet, as Skow notes, Brown rejects this: "Note that I do not assume that the metric is *defined* by the behaviour of rigid bodies" [p. 23]. What's more, Brown is careful to treat rods and clocks as complex entities that are governed by some dynamics, and not as privileged entities that could

play a role in the coordination aspect of Einstein's practical geometry.[10] Indeed, given Brown's emphasis on—and endorsement of—Einstein's misgivings about privileging rods and clocks in the development of special (and general) relativity, his work might even be taken as supportive of Poincaré's conventionalism. But of course this comparison can only be taken so far, as Brown advocates Minkowskian rather than Euclidean geometry, and speaks freely about the symmetries of dynamical laws without assuming some geometrical structure first. As explained above, Poincaré would not approve! What's more, a conventionalist approach like Poincaré's would require an argument in favor of Minkowski geometry on the basis of its simplicity relative to other geometrical structures. And even if such an argument could be made, it would get one nowhere in accounting for that geometry by appealing to the symmetries of dynamical laws, which is a central tenet of the dynamical approach.

Thus Brown's work cannot obviously be read as a version of either 'ism' under consideration. However, one way of construing the dynamical approach, which puts a particular emphasis on the set of possible coordinatizations of the material world, bears some resemblance to both. First, recall from above (p. 66) that the dynamical approach is not prohibited from assuming some kind of sub-metrical spatiotemporal structure, so long as the full structure of Minkowski spacetime is still somehow accounted for by the symmetries of the dynamical laws. And if only a sub-metrical structure were ascribed to material bodies, rigid bodies would clearly not exist. But consider a case in which the material world's sub-metrical spatiotemporal structure is such that, under certain coordinatizations related by the elements of some smaller symmetry group—i.e. smaller than that of the primitive sub-metrical spatiotemporal structure—there are particularly simple dynamical laws that can be expressed. The geometrical structure corresponding to that symmetry group could then be chosen as a convention, so that metrical structure becomes part of the package that is chosen on the basis of simplicity. Indeed, this would be a form of 'standpoint B', described above. However, according to the dynamical approach,

[10] See earlier comments in footnote 26 on page 59.

Minkowski geometry is not chosen for its own simplicity, or even for the simplicity of its pairing with some set of dynamical laws. Rather, Minkowski geometry is said to be picked out by the symmetries of the dynamical laws. And so for the dynamical approach to be understood a form of conventionalism, one must put a greater emphasis on the simplicity of the dynamical laws themselves, rather than the combination of geometry and dynamics. Thus it would be fairer to say that the dynamical approach might be understood as an adaptation of Poincaré's conventionalism, in which the criterion of simplicity is transferred from the choice of geometry to the choice of dynamical laws themselves, regardless of whether the corresponding geometrical structure (chosen by convention) is itself attractively simple. According to this form of conventionalism, it is recognized that a material system with sub-metrical geometrical structure can nonetheless be labeled by sets of coordinates that are related to one another by transformations corresponding to the symmetry group of a richer geometrical structure. When all sets of coordinatizations are compared, the set with respect to which the simplest dynamical laws can be written is used as the reference in choosing a conventional metrical structure.

But with all of that being said, the connection between the dynamical approach and any form of conventionalism can once again only be taken so far. For according to the dynamical approach, there is in fact no element of choice regarding geometrical structure. Rather, there is a matter of fact about geometrical structure—namely, that it is Minkowskian. How can the conventionalist approach just described be adapted to take this into account? One strategy would be to take a Humean, or 'non-governing' conception of laws. This conception of laws is introduced in more detail below (§4.1). For now, suffice it to say that according to this view, dynamical laws do not 'govern' the behavior of material objects—i.e., dynamical laws are not 'responsible' for their behavior in any way. Rather, laws are mere systematic descriptions of the patterns or regularities in the history of the material world.[11] This view is usually coupled with a 'best-systems' definition of laws, which is also introduced in more detail below (§4.1). According to

[11] Beebee [2000, §II] gives a helpful introduction to the non-governing conception of laws, contrasting it with the 'necessitarian' conception. See also Carroll's introduction to laws of nature [2011].

this view, the laws are theorems of the simplest and strongest such systematizations. By adopting a conception of laws along these lines, the element of choice can be removed from the adaptation of Poincaré's conventionalism, mentioned above. For instead of choosing, by convention, some geometry that supports particularly simple dynamics, it is the fact that some or other particularly simple dynamics *are the laws* that determines the corresponding geometrical structure. In this way, there can be a matter of fact about geometrical structure without it being taken as primitive—rather, it is determined by the best systematization.

Mattias Frisch suggested something similar in his review of Brown's work.

> [From] the perspective of a Humean best-system-account of laws together with spacetime relationalism it becomes even more difficult to see how the geometry of Minkowski spacetime could be explanatorily prior to the Lorentz invariance of the laws. Both the choices of geometrical structure to represent the motion of physical objects and the choice of dynamical laws, on this view, are determined by which choice yields the best system overall. The choice of metric, thus, is not prior to the determination of the laws but rather is made in conjunction with the choice of laws. If the simplest laws are Lorentz-invariant, then the simplest choice of metric is the Minkowski metric. Or, as Brown and Pooley put it: "the appropriate structure is Minkowski geometry *precisely because* the laws of physics, including those to be appealed to in the dynamical explanation of length contraction, are Lorentz covariant" [Brown and Pooley 2006, p. 80]. [Frisch 2011, p. 183]

Now, because Frisch speaks of a "best system overall", in which geometry is determined "in conjunction with" the laws, it would seem that he has in mind something along the lines of the third way out of the Einstein–Poincaré controversy, mentioned above (p. 75). But it was also explained above that given the explanatory claims of the dynamical approach, the real priority ought to be the simplicity of the laws themselves, and not on their pairing with geometry. According to the best-systems view, there is a matter of fact as to what those laws are. And so whereas Frisch continues to speak of some element of 'choice' when a 'Humean best-system-account' of laws is assumed, it should be stressed that this would no longer be a form of conventionalism in the pure sense: the 'choice' of

geometry would be fixed by what the best-system dynamical laws are.

In the end, this construal of the dynamical approach is in some ways more like Einstein's practical geometry than a form of conventionalism. For while no coordinations are established between allegedly rigid bodies and the elements of a 'pure' geometry, an emphasis is put on another kind of coordination: given the range of possible coordinatizations of the material world, those with respect to which the best-systems dynamical laws are expressed are themselves related by elements of the Poincaré group. In this way, Minkowski metrical structure is not taken as primitive, and so the arrow of explanation from dynamics to geometry is satisfied, and yet spatiotemporal structure is indeed a branch of science. In the following section, a similar idea is traced through some early work on special relativity. Then, the metaphysical details for this form of 'practical geometry', and the particular conception of dynamical laws that it involves, are outlined in more detail in Chapter 4.

3.2 A Trend in Minkowski, Schlick, and Brown

It has been argued above that the explanatory claims of the dynamical approach can be satisfied if the latter is understood as a form of practical geometry, according to which spacetime structure is determined by the dynamical laws written in terms of it. Sympathies for this kind of approach appear in some early works on special relativity, which are outlined below. Finally, Brown's own discussion on coordinatizing material events is shown to be consistent with this understanding of the dynamical approach.

Minkowski's Coordinatization Process

As might be suspected by the earlier quotations from Minkowski's work (§1.2), Minkowski's presentation on the geometrization of special relativity has at times been understood as a promoting the existence of a substantival 'container' spacetime, in which all material objects exist.[12] But a closer reading suggests otherwise, as Minkowski repeatedly refers to

[12]For example, Craig [2008, p. 13; 2001, pp. 77–8] takes this view, although he goes on to dispute there being any good reason to be a substantivalist [2001, p. 78, pp. 190–4]. See also §1.3, above.

the symmetry group of dynamical laws as the core of Einstein's theory. This is especially true by his emphasis on the 'fundamental world postulate', that all laws are Lorentz covariant. Brown has stressed this point more than once.[13]

> *Minkowski's paper is an attempt to spell out the mathematical consequences, only partially understood by Einstein, of the world-postulate, and in particular of the claim that the laws of physics are Lorentz covariant.*
>
> [Brown 2005*b*, p. 130]

But in fact a further argument can be made along the lines that Minkowski's presentation of the geometrization of special relativity is more consistent with the dynamical approach than the geometrical approach, especially when the dynamical approach is understood as suggested above. For in Minkowski's work there is also a brief discussion on how the Lorentz covariance of dynamical laws is to be discovered, which involves focusing first on material events *without* consideration of any geometrical structure, and then considering the range of possible coordinatizations of those events, and finally selecting the set with respect to which some preferred dynamics can be written.

Minkowski first introduces the idea of coordinatizations as "a system of values x, y, z, t", at each of which he takes there to be "something perceptible", and the whole collection of which he calls the "*world*" [1909, p. 76]. The only geometry initially ascribed to that world is whatever 'continuity' structure is suggested by his introduction of "*world-lines*", which he views as representative of the material world in its purest form.

> The whole universe is seen to resolve itself into similar world-lines, and I would fain anticipate myself by saying that in my opinion physical laws might find their most perfect expression as reciprocal relations between these world-lines.
>
> [Minkowski 1909, p. 76]

Next, after giving a brief mathematical introduction to the relativistic symmetry group G_c mentioned above (p. 18), Minkowski explains how we come to find that this is the symmetry group of our best dynamical laws.

> The existence of the invariance of natural laws for the relevant group G_c would have to be taken, then, in this way:—

[13] See also his second collaboration with Pooley [Brown and Pooley 2006, p. 37].

> From the totality of natural phenomena it is possible, by successively enhanced approximations, to derive more and more exactly a system of reference x, y, z, t, space and time, by means of which these phenomena then present themselves in agreement with definite laws. But when this is done, this system of reference is by no means unequivocally determined by the phenomena. *It is still possible to make any change in the system of reference that is in conformity with the transformations of the group G_c, and leave the expression of the laws of nature unaltered.* [Minkowski 1909, p. 79]

Now, as Dennis Dieks has recently pointed out, this paragraph reveals some sympathies for an approach to spacetime structure like the preceding construal of the dynamical approach. Minkowski begins by considering all possible coordinatizations of world-points, where these coordinatizations should arguably preserve the continuity of the aforementioned world-lines. He then selects some set of coordinatizations with respect to which particular dynamical laws can be written. That is, whereas Minkowski takes the laws to find their purest form in describing relations between world-lines, he takes them to find their most practical form in a particular set of coordinates related to one another by elements of the Poincaré group.[14]

> In a nutshell, Minkowski's suggestion is to do without any pre-given interpretation of coordinates in terms of distances and clock settings, and to start with a theoretical account of elementary physical phenomena in terms of completely arbitrarily chosen variables. The form of the regularities formulated in this way will most likely be extremely complicated, but by going to new independent variables, that is, by performing mathematical coordinate transformations, one may simplify. As the final result of this process of successive simplifications, a preferred set of space-time coordinates will be found as that "system of reference x, y, z, t, space and time, by means of which these phenomena then present themselves in agreement with definite laws" [Minkowski 1909]. [Dieks 2009, p. 239]

Of course, this is not to say that Minkowski's approach to spacetime structure was precisely the form of practical geometry advocated above.[15] But at the very least, it certainly

[14] Dieks goes too far in saying that Minkowski's coordinatization project involves "completely arbitrarily chosen variables." For as noted above, Minkowski does seem to presuppose some form of continuity structure on the world-points.

[15] And yet, Dieks goes so far as to compare Minkowski's approach with 'Leibnizian relationalism', according

strengthens Brown's argument that the starting point in Minkowski's presentation of the geometrical structure that now bears his name was the Lorentz covariance of dynamical laws. And as Dieks stresses in a related article, this is very different from the standard 'geometrical' approach.

> This approach is very different from the usual textbook approach, in which one begins with the symmetries of spacetime and then imposes these same symmetries on the physical laws. Minkowski's method does the opposite thing: one starts with physical systems and the regularities found in their behaviour, then formulates laws and studies the invariance properties of these laws and, finally, one calls the symmetries shared by all physical laws "spacetime" symmetries. [Dieks 2010, p. 237]

What's more, this also shows that Minkowski appreciated the importance of comparing possible coordinatizations in determining the dynamical symmetry group, which could then be used as a pointer to a corresponding geometrical structure. And this, of course, is a key element in the form of practical geometry that is being suggested as a way to understand the dynamical approach. Granted, this would not be the case if Minkowski is understood as assuming the world to have some primitive geometrical structure which is being revealed by his method of "enhanced approximations", for the dynamical approach is not compatible with taking Minkowski geometrical structure as primitive. But, if Minkowski is understood to assume only whatever continuity structure is entailed by world-points constituting continuous world-lines, then these enhanced approximations could be understood as a method of determining an appropriate geometrical structure that would support a relatively simple and useful physical description of the primitively less-structured material world.

---

to which less geometrical structure is taken as primitive than that suggested by the symmetries of the world's dynamical laws. Leibnizian relationalism is discussed briefly in §4.1, below, and plays a critical role in the construal of the dynamical approach being developed in this and the following chapter. Another discussion of Leibnizian relationalism, in a related context, is offered by Pooley [2013, §3.1].

Moritz Schlick's Philosophy of Space and Time

As a correspondent of Einstein's during the early 20th century, Moritz Schlick was influential in Einstein's own early interpretation of relativity.[16] Their correspondence began at least as early as 1915, when Schlick wrote a commentary (of which Einstein highly approved) on the special theory of relativity, *The Philosophical Significance of the Principle of Relativity* [1915].[17] In that work, Schlick goes into detail in presenting Einstein's principle of relativity as physically equivalent to the Lorentz–FitzGerald contraction hypothesis [§IV], and goes on to commend Einstein's principle as the simpler and thus preferred approach [§V]. But as Tom Oberdan has noted, Schlick's emphasis on simplicity is somewhat complex. For in comparing the aforementioned approaches to special relativity, Schlick noted differences in both the so-called 'representational scheme' and also in the formulation of empirical claims themselves.

> Implicit in the general philosophical scheme in which Schlick discussed Relativity is an objective, logical distinction between the representational framework in which scientific claims may be formulated and those claims themselves. An intrinsic function of the representational scheme is the constitution of the very concepts in which the formulation of empirical claims is first made possible. [Oberdan 2013, p. 10; *cf.* 2009, pp. 200–1]

In short, the Lorentz–FitzGerald hypothesis preserved a familiar representational scheme, but its descriptions of reality involved supplemental hypotheses. In contrast, Einstein's principle-theory approach used an allegedly more complicated (or at least less familiar) representational scheme, but enjoyed arguably simpler descriptions of reality.[18] Oberdan points out that in favoring Einstein's approach for that very reason, Schlick took a conventionalist view like the one outlined at the close of §3.1 above, according to which the criterion of simplicity was shifted away from the representational scheme of Euclidean

[16] Indeed, Einstein references Schlick for the idea of 'implicit' definitions of primitive terms in pure geometry [Einstein 1921, p. 234]. (See p. 70, above.) In the work that Einstein references [Schlick 1925, §7], Schlick cites Hilbert [1971].

[17] The Einstein–Schlick correspondence is also discussed by Howard [1984; 2014, p. 11], by Friedman [2002, fn. 39, p. 194, p. 203], by Giovanelli [2012, §3.1], by Oberdan [2013, p. 5, p. 15], and by Brown and Lehmkuhl [TBD, §4].

[18] Schlick makes similar comments elsewhere [1922, pp. 210–1].

geometry, away also from any combination of physics and geometry, and over to the dynamical laws themselves.[19]

> Parting from Poincaré, Schlick recognized that the representional framework that appears simplest when regarded in isolation may nonetheless require excessively complicated formulations for the description of reality. And he insisted—*contra* Poincaré—that it is the simplicity of these formulations that is the most compelling desideratum, not the simplicity of the representational scheme. Thus, the representational scheme that allows for the simplest description of reality is always to be preferred—so much the worse for Euclid, and Poincaré, too. [Oberdan 2013, p. 12]

Now, it should be granted that in referring to a 'representational framework', Schlick has something in mind more general than a mere set of potential spacetime structures. But of course when the conversation is restricted to mathematical models of spacetime, it is geometrical structure that plays the role of the representational framework. In a discussion of Schlick's work on general relativity, *Space and Time in Contemporary Physics* [1922], Oberdan makes the point in greater detail.[20]

> Following Poincaré, and Kant before him, Schlick recognized that geometry alone does not depict the object of physics—the motion of bodies—but only provides the representational scheme for the depiction of physical motion. Since the description of motion is the principal goal of physical theorizing, the criterion for the choice of geometry is the simplicity of the resulting description of motion, not the simplicity of the geometry itself. [...] Nor, for that matter, is the simplest overall formulation of the combination of physics and geometry the appropriate desideratum, as Einstein maintained. [Einstein 1921] Rather, it is the simplest expression of the theory's empirical content which is desirable, even when this requires complicating the representational scheme or, in the present case, the underlying geometry.
>
> [Oberdan 2009, pp. 203–4; *cf.* pp. 201–3]

Thus in Schlick's approach to relativity we find an illustration of the conventionalism outlined earlier, and to which the dynamical approach was briefly compared. And just as

[19] The same observation about Schlick's conventionalism is also made by Friedman [2002, p. 218].
[20] In addition to the references that Oberdan provides, it seems clear that his discussion draws from Schlick's comments earlier in the same work [1922, §§IV–V].

before, a Humean, best-systems conception of laws would make Schlick's conventionalism compatible with the explanatory claims of the dynamical approach. For if Schlick were not faced with a choice of geometry for the sake of simplifying the dynamical laws, but rather if dynamical laws were in fact the axioms of the simplest and strongest systematizations, then the corresponding geometrical structure would in turn be accounted for by the symmetries of those laws. In the case of a specially relativistic world, those laws would be Lorentz covariant, and the corresponding geometrical structure would be Minkowski spacetime.

Now, when something like this approach was uncovered in Minkowski's work, it was also shown that Minkowski seemed to take some form of continuity structure as primitive. Similarly, it turns out that Schlick was in favor of assuming some form of continuity structure for the material world. Given his later involvement in the Vienna Circle, it should not be surprising that he aimed to develop that structure empirically. In his work on general relativity, Schlick outlined a 'method of coincidences' based on observable coincidences, the end result of which was a topological structure, and which was all that Schlick took to be the fundamental geometrical structure of the world.[21] Indeed, in a comment not unlike Minkowski's above (p. 80), Schlick went so far as to suggest that this represented reality in its purest form.[22]

> If it were possible to formulate the laws of nature in such a way that they contained exclusively those propositions concerned with coincidences, which are true for *any* systems of reference whatsoever, they would then contain a minimum of arbitrary methods of description and would, in consequence, reproduce nature, so to speak, in the most faithful manner as possible.
>
> [Schlick 1949, p. 37]

Indeed, the fact that Schlick takes some topological—or perhaps differentiable—structure as primitive on the point coincidences is clear not only in his 'method of coincidences'

[21] This method of coincidences is addressed in at least two places in Schlick's work [1925, §31; 1949, Ch. 6]. And as Friedman points out [2002, pp. 204–5], it is in 1922 that Schlick shows how his method of coincidences ties in especially with general relativity. Oberdan also discusses Schlick's intention to assume only a topological structure to the world, in light of Einstein's work on general relativity [2013, pp. 20–1].

[22] And of course this is reminiscent of Einstein's well-known quote to the effect that "[all] our space-time verifications invariably amount to a determination of space-time coincidences" [1916, p. 117].

discussed in the works cited above, but also in comments scattered throughout the aforementioned work on general relativity [1922]. For instance, Schlick claims that anything which cannot be reduced to point-coincidences lacks objective value, and that any transformation on the coincidences or their coordinates that preserves the smoothness of coordinatization amounts to the same physical system.[23]

> The adjustment and reading of all measuring instruments of whatsoever variety [...] are always accomplished by observing the space-time-coincidence of two or more points. [...] Everything else in our world-picture which can *not* be reduced to such coincidences is devoid of physical objectivity, and may just as well be replaced by something else. All world-pictures which lead to the same laws for these point-coincidences are, from the point of view of physics, in every way equivalent. We saw earlier that it signifies no observable, physically real, change at all, if we imagine the whole world deformed in any arbitrary manner, provided that *after* the deformation the coordinates of every physical point are continuous, single-valued, but otherwise quite arbitrary, functions of its coordinates *before* the deformation (and the physical 'constants' behave accordingly). Now, such a point-transformation actually leaves all spatial coincidences totally unaffected; they are not changed by the distortion, however much all distances and positions may be altered by them.
>
> [Schlick 1922, pp. 240–1]

By determining a topological structure for those coincidences, Schlick provided a geometrical structure with respect to which all potential representational schemes must be consistent, so that it is only in terms of those structures which respect that topology that the dynamical laws could be written.

Whether Schlick takes spacetime to be a substantival entity, so that the geometrical structures mentioned above would be attributed to it, is a separate issue. But whatever his position on the substantivalist–relationalist debate, there is of course nothing to stop us from taking his strategy and applying it directly to material objects.[24] Thus in

[23] As mentioned by Friedman [1983, pp. 21–4], Schlick concludes earlier in the same work that space is "metrically amorphous" [1922, pp. 230–3]. Elsewhere, in a discussion on his 'method of coincidences', Schlick makes a similar comment [1925, p. 275]. Friedman also offers some discussion on this [2002, fn. 36], and the following quotation is also discussed by Oberdan [2009, pp. 204–5].

[24] Besides, even this form of substantivalism, in which the primitive geometrical structure is more symmetric than that chosen for the sake of simplifying dynamics, would be compatible with the dynamical approach. See footnote 31 on page 65, and the discussion in §6.1, below.

Schlick's approach to relativity, we see the aforementioned understanding of the dynamical approach foreshadowed. Schlick would have us take as primitive some minimal geometrical structure—so as to provide a continuity structure and coincidence relations—and then choose as a convention whatever representational scheme (i.e. geometrical structure) enables the simplest empirical descriptions (i.e. the simplest dynamical laws). But as suggested above and discussed further in the following chapter, taking on a Humean, best-systems conception of dynamical laws might allow the geometrical structure to be determined by the symmetries of the dynamical laws themselves.

Brown's "Abbreviated Constructive Project"

We turn now to Brown's own work which, if the above construal of the dynamical approach is to be definsible, ought to be compatible with the ideas introduced in this chapter. To see that it is, recall what Norton has called Brown's "abbreviated constructive project" [2008, p. 829]. In the opening of the second chapter of *Physical Relativity*, Brown outlines three questions that discerning students ought to ask when considering the principle of general covariance in Einstein's formulation of general relativity. The first of these occupies Brown's attention for the remainder of that section:

> It is common in discussions of the principle of general covariance in Einstein's general theory of relativity to find the claim that coordinates assigned to events are merely labels. [...] Before we consider labeling them, what *physically* distinguishes two different events of exactly the same kind?
>
> [Brown 2005*b*, p. 11]

Brown's response reveals that he has in mind an ontology of physical fields with some kind of sub-metrical structure. Such an ontology and ideology would be consistent with the form of practical geometry described above, and also with the more detailed illustration to be provided in Chapter 4, below.

Brown begins his response by casting doubt upon the existence of points of substantival spacetime. He reviews Robert Geroch's [1978] distinction between a localized 'material event'—like a firecracker or finger-snapping, and which Brown calls a physical 'marker'—

and the point of spacetime at which the event is taken to occur in a substantivalist perspective. And whereas Geroch was allegedly hesitant to make a distinction between the 'point' and its 'marker', Brown is clear in stating that he does not distinguish the two at all. That is, Brown is not interested in the relationships between points of spacetime, but rather in the relationships between their markers, which he defines as follows.

> Recall we can think of the markers as suitably idealized explosions, collisions of point-particles, flashes of light and so on—the kinds of things physicists typically mention when asked to provide examples of 'events'. We might even try to be more technical and insist on characterizing a marker as the set of values at a point of the (components of the) most fundamental fields in our best physical theories. [Brown 2005*b*, p. 13]

Here, Brown is most naturally read as advocating an ontological relationalism, even if it is not crucial to Brown's thesis. And in the place of substantival spacetime points, Brown suggests a fundamental ontology of 'markers' that comprise coincident values of physical fields.[25] Then, his ideology is revealed by his approach to distinguishing such markers for the purposes of coordinatization. As Brown explains, there will surely exist various types of markers that are intrinsically indiscernible.

> Whatever your favourite example of a marker is, it is bound to occur at many distinct points in space and at many times in the history of the universe. The very existence of a law-like structure in the universe, of the fact that physics deals with empirical *regularities*, makes this virtually inevitable. The flash of light, or the collision of particles or whatever, taken in its idealized, pristine, localized sense, is *not a one-off*. [...] As a consequence, in the language of the mathematical physicist, there simply cannot be a one-to-one correspondence between the set of distinct marks and the set of points that is the space-time manifold. [Brown 2005*b*, p. 13–4]

Here, it should be noted that Brown finds himself in a more difficult situation than the flat-footed relationalist. If Brown were to admit Minkowski spacetime structure as primitive, then indistinguishable markers would not be an issue for any but the most unrealistically

[25] This is consistent with what Brown has written elsewhere, that "space-time points are perhaps best viewed not as entities in their own right, but as correlations or links between the individual degrees of freedom of distinct physical fields" [1997, p. 68]. Pooley makes a similar observation [2014, fn. 48, §5.8].

homogeneous or symmetric universes: the possible coordinatizations would be fixed by the requirement that they respect the structural properties that supervene upon the field values' spatiotemporal properties.[26] But so far, Brown has only admitted markers, many of which will arguably be intrinsically indistinguishable.[27] In order to justify some choice of coordinatization, Brown must also admit an appropriately rich spatiotemporal structure without undermining the goal of accounting for metrical structure in terms of the symmetries of dynamical laws.

Toward that end, Brown argues that each individual event stands in a unique relationship to the universe as a whole. That is, events "have different coordinates because they are distinct, and they are distinct not in virtue of what they are *locally* but in virtue of the fact that *they stand in different relations to the rest of the universe*—to the rest of the markers" [p. 14].[28]

> Let's consider what it is that distinguishes two [events that] do not coincide. The distinction does not lie in the fact that they have different coordinates. They have different coordinates because they are distinct, and they are distinct not in virtue of what they are *locally* but in virtue of the fact that *they stand in different relations to the rest of the universe*—to the rest of the [events]. It is because those relations are in principle discernible that we can say that the same [events] can occur at different space-time points. So rather than think of a space-time point [...] as a self-contained localized atom of the invisible uniform space-time manifold, we might more usefully think of it as the view of the universe from a point. [Brown 2005*b*, p. 14]

To illustrate, Brown borrows a phrase from Julian Barbour [1982].

> Minkowski, Einstein, and Weyl invite us to take a microscopic look, as it were, for little featureless grains of sand, which, closely packed, make up space-time.

[26] Pooley offers further discussion [2014, §5.8].

[27] I take Brown to use the term 'marker' to refer to a material event 'type', and to use the term 'event' to refer to a material event 'token'.

[28] Here, Brown might be read as making the particularly strong claim that two instantiations of some marker are distinct *in virtue of* differences in their spatiotemporal properties—i.e. that their distinctness supervenes upon spatiotemporal structure. However, I take Brown to be making the weaker claim: simply that in worlds like ours, two distinct events are so discernible. Of course, this would entail a tacit assumption that two events can in principle be distinct without differing along the lines described above. That is, Brown appears to rely upon some notion of 'weak discernibility', as has been advocated by Simon Saunders [2003; 2006]. Pooley discusses Saunders's position and related views in context [2014, §3.3].

> But Leibniz and Mach suggest that if we want to get a true idea of what a point of space-time is like we should look *outward* at the universe, not inward into some supposed amorphous treacle called the space-time manifold. The complete notion of a point of space-time in fact consists of *the appearance of the entire universe as seen from that point*. Copernicus did not convince people that the earth was moving by getting them to examine the earth but rather the heavens. Similarly, the reality of different points of space-time rests ultimately on the existence of different (coherently related) viewpoints of the universe as a whole. Modern theoretical physics will have us believe that the points of space are uniform and featureless; in reality, they are incredibly varied, as varied as the universe itself. [Brown 2005*b*, p. 14]

Of course, Barbour's suggestion is easy enough to understand when the world is taken to have some kind of Riemannian, pseudo-metrical structure. But what sense can be made of this if some such structure cannot be assumed, so that it remains to be accounted for by the symmetries of the dynamical laws? Brown has not yet ascribed any specific sub-metrical structure to the world's collection of markers, and so it is not clear what 'view' of each other the events are supposed to have, nor in virtue of what relational properties this 'view' arises. Granted, questions of this sort aren't Brown's primary focus, but it is clear that Brown must be assuming at least *some* kind of primitive ordering or arrangement to the world's markers. Indeed, Brown's markers apparently have all the sub-metrical continuity structure of a differentiable manifold, but allegedly not any metrical structure. Thus it would be fair to say that Brown's fundamental metaphysics seems to be a collection of markers instantiating some particular *topological*, or perhaps even *differentiable* structure. And in fact, Norton cites personal correspondence that supports this.[29]

> In the process of the refereeing of an earlier version of this note, Harvey Brown has assured me that [Brown and Pooley 2006] and [Brown 2005*b*, especially Ch. 2] presumed the existence of a manifold of spacetime events with coordinate systems. [Norton 2008, p. 829]

Now, whether this kind of primitive structure would justify the idea that each event

[29] Nick Huggett also makes a similar observation [2009, pp. 417–8].

has a *unique* 'view of the universe', which could be used to distinguish it from all other events, can for present purposes be left aside.[30] More pertinent to the project at hand is to review what remains to be worked out in Brown's metaphysical picture. In particular, where do Brown's explanatory, Lorentz-covariant dynamical laws come into the picture, and how are we to make sense of Lorentz-covariance if not in terms of coordinatizations that tacitly assume some spatiotemporal structure that supersedes the 'rubber sheet' structure of a topological or differentiable manifold? There is a need for some further argument, to the effect that there is *something about* the world's topologically or differentiably structured material events such that, given all systematizations made possible by all possible coordinatizations, some particular Lorentz-covariant equations are to be preferred. Only then would one have a metaphysical picture that does not presuppose the structural properties of Minkowski spacetime, and from which it could be said that the dynamical laws account for Minkowski spacetime structure. It has been argued above that a particular Humean conception of dynamical laws plays a crucial role in such a project. In the following chapter, that conception of laws is outlined in more detail, as is the idea of taking only some sub-metrical structure as the primitive geometry of the material world.[31]

[30] But see the related discussion in footnote 11 on page 104, regarding the 'discriminating' aspect of topological nearness relations.

[31] Despite the fact that Brown sometimes quotes his work approvingly, the interpretation of the dynamical approach defended here should not be confused with the more empiricist approach promoted by DiSalle [1995], which is discussed in a similar context by Pooley [2014, Ch. 6].

4 Supervenient Spacetime:
The Dynamical Approach as Regularity Relationalism

In order to take seriously the construal of the dynamical approach presented in the previous chapter, a first step will be to outline the details of how a Humean conception of laws could allow dynamical symmetries to explain or account for the very geometrical structure in terms of which they are written. Such ideas are at the heart of Nick Huggett's recent work on 'regularity relationalism', which is introduced in §4.1. A Humean conception of natural laws features prominently in regularity relationalism, but just as in Lewis's own description of Humean supervenience, introduced below, Huggett's regularity relationalism also includes metrical relations in the fundamental ideology. For that reason, Humean supervenience and regularity relationalism may seem seem inhospitable to the idea of choosing topology as the primitive geometrical structure of one's 'Humean mosaic', as suggested in Chapter 3.[1] However, in §4.1 it is also shown that topological structure is quite compatible with the core tenets of Humean supervenience, and that taking topology as primitive can actually resolve certain problems that Lewis faces in his closely related work on modal realism.

Understanding the dynamical approach as a form of regularity relationalism that takes topological structure as primitive has recently been suggested by Oliver Pooley, whose work is introduced in §4.2. There, some discussion is given as to what particular topological structure would best suit the dynamical approach. Reasons are given for choosing the standard Euclidean topology of $\mathbb{R}^4$, which is introduced briefly, and which is referred to as $\mathcal{R}^4$. With that choice settled, §4.3 rehearses and illustrates the 'constructivist' project of the dynamical approach in a more precise way, and this sets the stage for a case-study to be undertaken in Chapter 5.

[1] As mentioned above, a brief introduction to point-set topology, and its place at the bottom of the standard hierarchy of spacetime geometrical structures, is provided in Appendix A.

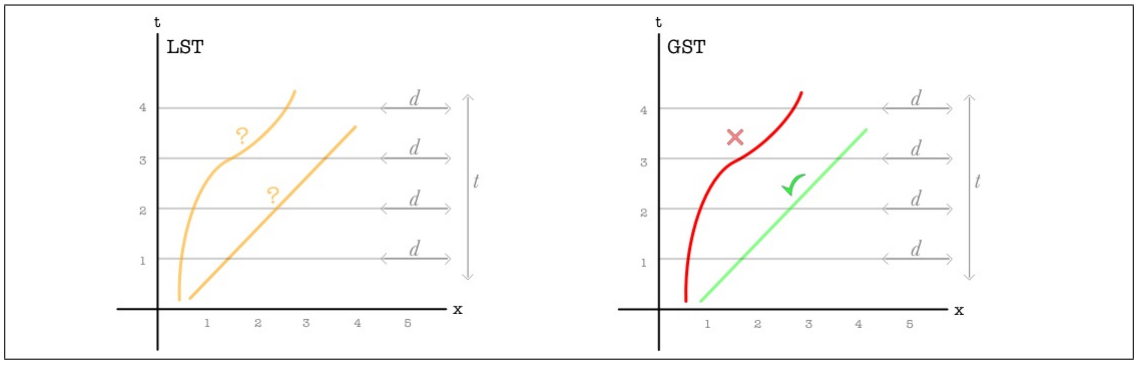


Figure 3: Following Earman's terminology [1989, Ch. 2], Galilean spacetime has the same spatial (d) and temporal (t) metrical structure as Leibnizian spacetime. But in addition, it also has an affine structure.

4.1 Regularity Relationalism and Primitive Topological Structure

Nick Huggett has recently promoted 'regularity relationalism' as a way to overcome the problems faced by a Leibnizian relationalist in a Newtonian world [2006]. Leibnizian relationalism assumes a material ontology, and the spatiotemporal components of its ideology are only those relations encoded by Leibnizian spacetime: a 'simultaneity' equivalence relation; a temporal metrical relation on the space of simultaneity slices; and a spatial metrical relation on pairs of points in each simultaneity slice.[2] The trouble for Leibnizian relationalism is that this amounts to only part of the apparent geometrical structure of a world in which Newton's laws hold. Instead, the full geometrical structural presupposed by Newton's laws is captured in Galilean (or Neo-Newtonian) spacetime.[3] As illustrated in Figure 3, the latter is more structured—i.e. less symmetric—than Leibnizian spacetime. And for present purposes, the salient difference is Galilean spacetime's affine, or inertial, structure. A world in which the law $F = ma$ holds is a world in which there is a matter of fact about whether a material object is accelerating, and is therefore a world in which there is a matter of fact about whether a material object's worldline is curved or straight. By not underwriting any affine structure, Leibnizian relations do not account for the full geometrical structure assumed by Newton's laws.

[2]The presentation of Leibnizian and Galilean spacetime symmetries draws on Earman's [1989, §§2.2–2.4].
[3]Or, if gravity is taken into account, the apparent geometrical structure is captured by Newton–Cartan spacetime. See also footnote 27 on page 16. Newton–Cartan theory is outlined in detail by Malament [2012, §4]. For a related discussion, see recent work by Knox [2013], as well as recent and forthcoming work by Pooley [2013, §6.1.1; 2014, §4.6].

A more formal way of expressing this problem comes in terms of the symmetry groups of Leibnizian and Galilean spacetimes. Galilean spacetime is symmetric under translations, boosts, and static rotations, but Leibnizian spacetime is symmetric under those same transformations as well as accelerations and time-dependent rotations.

$$\mathcal{S}_{\text{G}} : \bar{r}\,' = \mathbf{R}\bar{r} + t\bar{v} + \bar{c} \quad ; \quad t' = t + d$$

$$\mathcal{S}_{\text{L}} : \bar{r}\,' = \mathbf{R}(t)\bar{r} + \bar{a}(t) \quad ; \quad t' = t + d$$

Thus in order to maintain that some world in which Newton's laws hold has less spatiotemporal structure than that of Galilean spacetime, the Leibnizian relationalist must give an account for the fact that the symmetry group $\mathcal{S}_{\text{L}}$ of the material objects' primitive spatiotemporal structure is actually a *supergroup* of that of the dynamical laws, $\mathcal{S}_{\text{G}}$. It might be argued that the relationalist's only forthright and honest response would be to posit a more robust ideology, which would result in fewer spatiotemporal symmetries, or else to propose some new, empirically adequate dynamical laws with a larger symmetry group, which would be written in terms of less spacetime structure.[4] But rather than take either of these approaches, Huggett proposes a move that he calls 'regularity relationalism'.

Regularity Relationalism

Regularity relationalism is effectively a way to have one's cake and eat it too.[5] The idea is simply that the Leibnizian relationalist can refuse to budge on taking only Leibnizian relations as primitive, by somehow explaining away the structural deficiency introduced above. To do so requires that one adopt a form of Humean supervenience, along with a 'liberalized' best-system conception of dynamical laws.

As for the former, consider the following summary by David Lewis.

[4] Moves like this are evaluated by Maudlin [1993] and Pooley [2013, 2014].

[5] Huggett also hints at this kind of approach in an earlier publication [1999, pp. 22–3]. The approach is referred to as 'have-it-all' relationalism by Pooley [2013, §6.3; 2014, §6]. Less favorable descriptions come from Belot [2011, §III], Arntzenius [2012, §5.7], and Dorr [2010, §VI; 2011, pp. 146–7], some of which are discussed in §6.1, below.

> Humean supervenience is named in honor of the greater denier of necessary connections. It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. [...] We have geometry: a system of external relations of spatiotemporal distances between points. Maybe points of spacetime itself, maybe point-sized bits of matter or aether or fields, maybe both. And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that.
>
> [Lewis 1986*b*, pp. ix–x]

As Lewis makes clear, the fundamental ontology of Humean supervenience amounts to the point-like parts of spacetime and/or physical fields, and the spatiotemporal component of its fundamental ideology includes some geometrical structure. As a form of relationalism, Huggett's regularity relationalist would naturally strike "points of spacetime itself" from the ontology, leaving the 'Humean mosaic' to comprise only "point-sized bits of matter". But the more interesting and salient characteristic of regularity relationalism comes up in its fundamental ideology. For according to Huggett's form of Humean supervenience, the full geometrical structure required to make sense of the world's dynamical laws is not taken as a primitive feature of the Humean mosaic. That is, in Lewis's quotation above, the regularity relationalist would take "geometry", "distances", and "arrangement" to mean only those constituted by a history of Leibnizian relations, which do not underwrite any affine structure.

Then, the 'liberalized' best-systems conception of dynamical laws is a version of the Mill-Ramsey-Lewis approach, according to which dynamical laws are theorems of the 'best' axiomatizations of the Humean mosaic [Lewis 1973, pp. 72–7]. The basic view can be liberalized by taking into consideration those laws that are written in terms of more geometrical structure than appears in the Humean mosaic. If the equations of some such systematization have a better balance of simplicity and strength than those written in terms of the mosaic's primitive geometry, then they are awarded the status of

dynamical laws. Whatever extra geometrical structure plays a role in the expression of those dynamical laws is simply claimed to be part of the "all else" that supervenes upon the mosaic's arrangement of qualities. Huggett describes the situation as follows.

> The admission of dynamical, supervenient quantities is what is new to this 'Humean' account. The MRL [Mill-Ramsey-Lewis] account admits only supervenient laws formulated in terms of natural properties. My proposal is that there are a wider range of strategies that can be employed in the service of systematizing a Humean (in this case, relational) history; the strongest-simplest system might involve laws formulated in terms of natural properties and supervenient properties. Thus the regularity account is strictly at odds with MRL; however it is certainly within its Humean spirit, simply utilizing a new strategy for systematizing the natural facts. [Huggett 2006, p. 50]

Thus regularity relationalism differs from the traditional relationalist approach to spacetime structure insofar as the former is a style of relationalism "according to which spatiotemporal facts are determined, not by relations at a time, but by regularities in the entire history of relations" [2006, p. 42].[6] And in the context of Newtonian mechanics, regularity relationalism amounts to the claim that the equations that best systematize the world's Leibnizian relational history are Galilean covariant ones. Thus Galilean spacetime structure supervenes upon the arrangement of qualities, and the dynamical laws are expressed in terms of it.

In effect, Huggett is taking an approach to Newton's laws that is reminiscent of van Fraassen's, according to which inertial frames are denied any privileged status. Instead, van Fraassen argues that there merely happen to be frames in which Newton's laws hold, but that this does not make those frames privileged in any way.

> The laws of motion are a set of statements about mass, motion, and force; therefore, they will be true in some frames of reference, in none, or in all. As it happens, they are true in some frames; these we call the inertial frames.

[6] It should be noted that Huggett needn't make a contrast between "relations at a time" and "regularities in the *entire history* of relations" (my emphasis). Whatever laws hold for the entire history of relations also hold in the local mosaic, and so it could also be said that regularity relationalism differs from the traditional relationalist approach to spacetime structure insofar as the former is a style of relationalism according to which spatiotemporal facts are determined by regularities across a finite section of the mosaic.

> [... But] the fact that the laws of a given theory hold only in some frames of reference can, as such, imply nothing about the status of these frames in nature. [van Fraassen 1970, p. 116]

Huggett, however, thinks in terms of an entire relational history—i.e. a Humean mosaic—and allows the dynamical laws to be written in terms of more geometrical structure than ascribed primitively to that relational history.

> I claim first that Newton's laws formulated as van Fraassen proposes plus a suitable force law constitute the simplest, strongest, systematization of the [Leibnizian] relational history; that they are indeed the (MRL) laws of the actual relational world. Second, since these laws supervene on the relational history and since they pick out the inertial frames, I claim that the inertial frames (and hence absolute accelerations) supervene upon the history of relations: inertial frames are the frames in which the laws that supervene on the history of relations hold; absolute acceleration is acceleration in the frames in which the laws that supervene on the history of relations hold.
>
> [Huggett 2006, p. 48, *cf.* pp. 46–7]

And so it is that the Leibnizian relationalist can allegedly account for the affine structure of a Newtonian world by admitting no more spatiotemporal structure to the Humean mosaic than that constituted by a Leibnizian relational history, so long as the 'liberalized' best systematization of that relational history happens to be Newton's laws. The catch, again, is that regularity relationalism takes the best systematization of the world to be one that is written in terms of some structure that does not exist primitively, but rather that supervenes upon those very laws' being best.[7]

However controversial it may be, regularity relationalism is as helpful for an advocate of the dynamical approach to special relativity as it is for the Leibnizian relationalist

[7] In this approach, the symmetry group of the laws is a subgroup of that of the primitive spacetime structure. This might be seen as a violation of Earman's second symmetry principle (p. 24, above). And given the comments a the close of §1.3, to the effect that the dynamical approach does not violate Earman's symmetry principles in the way that Craig's neo-Lorentzian approach does (fn. 54, p. 32), this might seem to go against the spirit of the dynamical approach. But note that the earlier comments were being made in reference to the fact that Craig's neo-Lorentzian view postulates spatiotemporal structure above and beyond that required to express the Lorentz-covariant dynamical laws. The dynamical approach, on the other hand, involves a violation of the other of Earman's symmetry principles, since it provides an argument to the effect that *not all* of the spatiotemporal structure in terms of which the dynamical laws are written must be taken as primitive. And of course, dynamical symmetry group would indeed match the spacetime symmetry group when supervenient geometrical structure is taken into account.

in a Newtonian world. For note that the Leibnizian relationalist's troubling position is precisely that of the dynamical approach, as described in Chapter 3—namely, the position of having admitted less primitive geometrical structure than that in terms of which the dynamics is written, so that there is some work to be done in accounting for the difference. And while it should be granted that Brown does not directly promote a Humean conception of laws in any of his work on the dynamical approach, it was argued in Chapter 3 that a Humean conception of dynamical laws would support the explanatory claims of the dynamical approach. What's more, just as Huggett draws upon van Fraassen's formulation of Newton's laws, Brown gives a similar account of Newton's first law.

> Newton's first law of motion [...] can be construed as an existence claim. Inertial coordinate systems are those special coordinate systems relative to which the above conspiracy, involving rectilinear uniform motions, unfolds. *Qua* definition, this statement of course has no content. What *has* content is the claim that such a coordinate system exists, applicable to *all* the free bodies of the universe. [Brown 2005*b*, p. 15]

Indeed, the idea that all possible coordinatizations should be included in determining the best-system dynamical laws was at the heart of the construal of the dynamical approach presented in the previous chapter. After exploring the possibility of adapting Huggett's regularity relationalism to take as primitive only some topological structure, the dynamical approach is introduced as a version of relativistic regularity relationalism that accounts for *Minkowski* spacetime structure in §4.2.

Primitive Topological Structure for Humean Supervenience

Huggett's regularity relationalism may provide the conception of natural laws necessary to understand the dynamical approach along the lines proposed in Chapter 3, but of course there are clear differences between the ontology and ideology that Huggett assumes in accounting for the affine structure of a Newtonian world, and that which was suggested above as being natural to take as primitive when accounting for the Minkowski

geometrical structure of a specially relativistic world. For whereas Huggett assumes an ontology of point particles and an ideology the spatiotemporal component of which comprises Leibnizian relations, it was shown at the end of §3.2 that Brown seems to assume an ontology of material events and an ideology the spatiotemporal component of which amounts to some sub-metrical structure—arguably a topological or differentiable structure. And while various reasons might be given for assuming *at least* a topological structure in providing an account for Minkowski spacetime structure, this seems to be in tension with regularity relationalism, which takes certain metrical relations as primitive. But in fact, taking some topological structure as primitive is not only compatible with Humean supervenience, it can also resolve certain issues acknowledged by Lewis himself.[8]

To begin, recall how Lewis summarizes Humean Supervenience (p. 96, above). Of particular importance here is the claim that "We have geometry: a system of external relations of spatiotemporal distances between points." In later work, Lewis clarifies the kind of relations that he has in mind, and their centrality in Humean Supervenience.

> Humean supervenience is yet another speculative addition to the thesis that truth supervenes on being. It says that in a world like ours, the fundamental relations are exactly the spatiotemporal relations: distance relations, both spacelike and timelike, and perhaps also occupancy relations between point-sized things and space time points. And it says that in a world like ours, the fundamental properties are local qualities: perfectly natural intrinsic properties of points, or of point-sized occupants of points. Therefore it says that all else supervenes on the spatiotemporal arrangement of local qualities throughout all of history, past and present and future. [Lewis 1994, p. 474]

And so it is clear that spatiotemporal (interval) relations are present in Lewis's 'Humean mosaic'. But as explained toward the end of Chapter 2 above, this wouldn't do for the dynamical approach. Fortunately, there is no principled reason why spatiotemporal relations must be taken as primitive in order that "all else supervenes on the spatiotemporal arrangement of local qualities". That is, unless an example can be given of some desired property that could not otherwise supervene, it would seem that any geometrical structure

[8]Thanks are due to John Dougherty for pointing this out.

is in principle a candidate for the spatiotemporal structure of Lewis's mosaic. And indeed, Lewis presents the above description of Humean supervenience as being adaptable in certain ways. The quotation above continues:

> The picture is inspired by classical physics. [...] The point of defending Humean Supervenience is not to support reactionary physics, but rather to resist philosophical arguments that there are more things in heaven and earth than physics has dreamt of. Therefore if I defend the *philosophical* tenability of Humean Supervenience, that defence can doubtless be adapted to whatever better supervenience thesis may emerge from better physics.
>
> [Lewis 1994, p. 474]

And if the natural/primitive local properties can be adapted to whatever better supervenience thesis may emerge from better physics, then surely so could the natural/primitive relations be adapted to the supervenience thesis that has emerged from the present exploration of the dynamical approach.

Another reason that spatiotemporal interval relations are not a critical component of Humean Supervenience comes up when the latter is tied in with Lewis's larger metaphysical project in *On the Plurality of Worlds* [1986*a*]. There, Lewis defends modal realism: "the thesis that the world we are part of is but one of a plurality of worlds, and that we who inhabit this world are only few out of all the inhabitants of all the worlds" [p. vii]. All such worlds are assumed to comprise the same fundamental ontology (spacetime points are the 'particulars'), and for the most part the same fundamental ideology (the spacetime points' purely natural properties). But how then are these worlds individuated? By the fact that the particulars do not share the same purely natural relations. That is, it is the lack of spatiotemporal relatedness that individuates one world from another.

> What, then, is the difference between a sum of possible individuals that is a possible world, and one that is not? What makes two things worldmates? How are the worlds demarcated one from another? [...]
>
> I gave part of the answer [...] when I said that nothing is so far away from us in space, or so far in the past or the future, as not to be part of the same world as ourselves. The point seems uncontroversial, and it seems open to

> generalisation: whenever two possible individuals are spatiotemporally related, they are worldmates. If there is any distance between them — be it great or small, spatial or temporal — they are parts of one single world. [...]
>
> So we have a sufficient condition: if two things are spatiotemporally related, then they are worldmates. The converse is much more problematic. Yet that is more or less the doctrine that I propose. Putting the two halves together: things are worldmates iff they are spatiotemporally related. A world is unified, then, by the spatiotemporal interrelation of its parts.
>
> [Lewis 1986*a*, pp. 70–1]

From this, we see that Lewis's Humean mosaic includes spatiotemporal relations not because they are necessary in order that "all else supervenes on the spatiotemporal arrangement of local qualities", but only because they are what individuate one world from another in Lewis's modal realism. But as it happens, using spatiotemporal relations in this way gives rise to some problems, which Lewis acknowledges.

In *On the Plurality of Worlds*, Lewis points out that relying upon spatiotemporal relations to unify the members of possible worlds would not allow for the possibility of a Newtonian world, in which there are no spatiotemporal relations, but rather spatial relations and temporal relations.

> Imagine a theory of spacetime that is built for Newtonian mechanics, or for common sense. [...] This theory will say that any two spacetime points are related by a spatial distance and a temporal distance: two different distances. [...] I suppose this is a way the world might have been, therefore it is a way that some world is. [...]
>
> We name the properties and relations that figure in our world; so what we call 'spatiotemporal relations' are relations that behave in the relativistic way, with spatial or temporal distance but not both. Now when we talk about the Newtonian world, are we talking about the possibility of different behaviour on the part of those same relations? Is it that those very relations might double up to give us two distances, one of each kind, between the same two points? Or are we talking instead about some different relations that might take the place of the spatiotemporal relations of our world?
>
> If it is the former, no worries. The Newtonian world is just as much spatiotemporally interrelated as ours is, even if the spatiotemporal relations behave differently there. But if it is the latter, then strictly speaking I cannot

> say that the Newtonian world is *spatiotemporally* interrelated.
>
> [Lewis 1986*a*, pp. 74–6]

Lewis admits that this problem seems to "call for a retreat", and acknowledges that a "last resort" would be to assume some sort of primitive worldmate relation [p. 74]. But instead, he outlines the criteria for some alternative, purely natural relation that might do, in other worlds, the job of unification that spatiotemporal relations do in ours.

> What I need to say is that each world is interrelated (and is maximal with respect to such interrelation) by a system of relations which, if they are not the spatiotemporal relations rightly so called, are at any rate analogous to them. Then my task is to spell out the analogy. At least some of the points of analogy should go as follows. (1) The relations are *natural* [...]. (2) They are *pervasive* [...]. (3) They are *discriminating* [...]. (4) They are *external* [...]. When a system of relations is analogous to the spatiotemporal relations, strictly so called, let me call them *analogically* spatiotemporal.
>
> [Lewis 1986*a*, pp. 75–6]

The details of Lewis's four characteristics are outlined in more detail below. In recent work, John Dougherty [2013] has pointed out that a certain 'nearness' relation (call it δ) would satisfy Lewis's criteria.[9] What's more, it can be formalized in such a way that it can, by satisfying certain axioms, generate a topology. Because some of the ways by which the relation satisfies the criteria for being analogically spatiotemporal hinge upon details of the axioms that it must satisfy in order to instantiate a topology, the latter should be addressed first.

That a nearness relation can establish a topological structure is relatively straightforward. Dougherty outlines a demonstration from other sources [Joshi 1983, p. 214; Cameron *et al.* 1974], which goes roughly as follows. Let $x \,\delta\, A$ mean that x is near set A. Let the domain be X, so that if $x \,\delta\, A$, then $x \in X$ and $A \subset X$. Then, $\langle X, \delta \rangle$ is a topological space whenever the following four axioms are satisfied.

1. Null: $\forall x (x \not\delta \emptyset)$

[9] Dougherty's work focuses primarily on using some such nearness relation as a predicate of resemblance for Lewis's work on a theory of universals [Lewis 1983].

2. Extensive: $\forall x(x \in A \rightarrow x\,\delta\,A)$
3. Union: $x\,\delta\,(A \cup B) \leftrightarrow (x\,\delta\,A) \vee (x\,\delta\,B)$
4. Idempotent: $(x\,\delta\,A) \wedge \forall a(a \in A \rightarrow a\,\delta\,B) \rightarrow x\,\delta\,B$

Indeed, it can also be shown that there is a one-one correspondence between the possible topologies on a given set and the possible instantiations of nearness relations on that set [Joshi 1983, p. 114]. And so in order for a world to be endowed with any particular desired topological structure, only the corresponding δ need be taken as a primitive relation.

As for being analogically spatiotemporal, it should be stated from the outset that in some sense, a 'nearness' relation is clearly analogically spatiotemporal—especially if it can indeed instantiate a topology. But to be sure, recall Lewis's criteria listed above. First, the relation must be 'natural'. This is easily satisfied, simply on account of the fact that these nearness relations are being proposed as 'primitive' relations, meaning that (in Lewis's terms) they are being taken as the 'perfectly natural relations'.[10] Second, the relation must be 'pervasive', which Lewis takes to mean that "mostly, or perhaps without exception, when there is a chain of relations in the system running from one thing to another, then also there is a direct relation" [p. 76]. Indeed, the axiom of idempotence ensures pervasiveness. Third, the relation must be 'discriminating', meaning that no two of the elements in the space should share all the same nearness relations. This criterion is satisfied by the fact that the points of some such topological space can indeed be individuated by the sum of their nearness relations, so long as the resulting structure satisfies the most basic separation axiom of topological spaces.[11] And finally, the relation must be external, meaning that the relations "do not supervene on the intrinsic natures of the *relata* taken separately" [p. 76, *cf.* p. 62]. Of course, that the nearness relation not supervene in this way could be stipulated by taking the nearness relation as primitive. Or, to ensure that it not be interpreted as supervening, we need only require a discriminating

[10] For more on Lewis's 'natural' relations, see the introduction by Hall [2012, p. 10]. A recent review of both enthusiastic and skeptical responses to the notion of naturalness is given by Dorr and Hawthorne [2013].

[11] Namely, the T_0 axiom, which is introduced in Appendix A and discussed at greater length in Appendix B. By definition, the points of a T_0 topological space are at least weakly discernible.

nearness relation on a set of material events of which at least one pair have identical intrinsic values (field values, etc.). Such a set would include some intrinsically indiscernible material events, which would nonetheless be discernible in virtue of the (discriminating) nearness relation

And so it is that a single 'nearness' relation, if taken as primitive and stipulated to satisfy the axioms above, can instantiate a topological structure. What's more, by the fact that δ can serve as an analogically spatiotemporal relation for Lewis's modal realism, this makes topological structure a suitable choice for one's Humean mosaic. In fact, rather than serving as an analogically spatiotemporal relation for only those aforementioned Newtonian worlds, it might also be taken as primitive in our world, or indeed in all worlds. This would amount to what Lewis calls the "nice" possibility that all worlds be individuated by one and the same (analogical) spatiotemporal relation [1986*a*, p. 75], rather than by different ones, as he seemed to suppose would be necessary in the quotation above. Doing so would require, of course, that the dynamical laws be expressed in terms of topological structure—or in terms of the nearness relation itself. Alternatively, with a liberalized version of best-systems dynamical laws, those laws could be expressed in terms of some supervenient spacetime structure, while the nearness relation serves as the analogically spatiotemporal relation unifying and distinguishing all possible worlds.

In closing, it should be noted that while the nearness relation δ is useful in illustrating the concepts of topological structure, it is certainly not the only primitive that can be axiomatized to describe a topological space. In fact, the points made above about how δ could serve as Lewis's analogically spatiotemporal relation could also be made with respect to other primitive relations, such as the 'open set' relation that appears in standard introductions to point-set topology (see Appendix A). Indeed, in what follows, topological structure is discussed in terms of its open-set structure, rather than in terms of a primitive 'nearness' relation. And in the final chapter (§6.1) the metaphysical details of taking topological structure as primitive are also discussed in terms of open sets.

4.2 Relativistic Regularity Relationalism

Above, it was suggested that the dynamical approach to relativity might be understood as a form of regularity relationalism for Minkowski spacetime. According to this construal, some topological structure is taken as a primitive feature of the material world, while Minkowski geometry is shown to supervene upon the symmetries of that world's 'liberalized' best-system dynamical laws. Here, it should be recalled that the inspiration for such a move is primarily to provide some metaphysical undergirding for the explanatory claims of the dynamical approach, especially in response to recent criticisms like those of Norton [2008], introduced above (§2.3). Interestingly, Norton acknowledges the possibility of taking a topologically structured spacetime as primitive [2008, p. 833], although no comments are made as to whether that topological structure could instead be assigned to the material objects themselves. A similar suggestion has also been made by Nick Huggett in his review of Brown's book.

> One possibility is that unlike geometry, topology is fundamental [... The] topology of the fields plus point coincidences do not determine the metrical properties, so there is additional work to be done by the laws, along the lines sketched above. [Huggett 2009, p. 418]

And more recently, Oliver Pooley has suggested something along these lines [2013, §6.3.2; 2014, §6.2]. First, Pooley explains that the dynamical approach should be construed as a relativistic version of Huggett's regularity relationalism.

> Recall that Huggett's regularity relationalist postulates primitive Leibnizian relations but no ideology corresponding to inertial structure. The latter is grounded in the existence of a proper subset of the coordinate systems adapted to the Leibnizian relations with respect to which the description of the entire relational history is the solution of particularly simple equations (Newton's laws expressed with respect to inertial frame coordinates). The dynamical approach involves a similar but much more radical move: the [Minkowski] metrical relations themselves are to be grounded in exactly the same way. [Pooley 2013, pp. 571–2]

But whereas Huggett's regularity relationalist takes as primitive the spatiotemporal

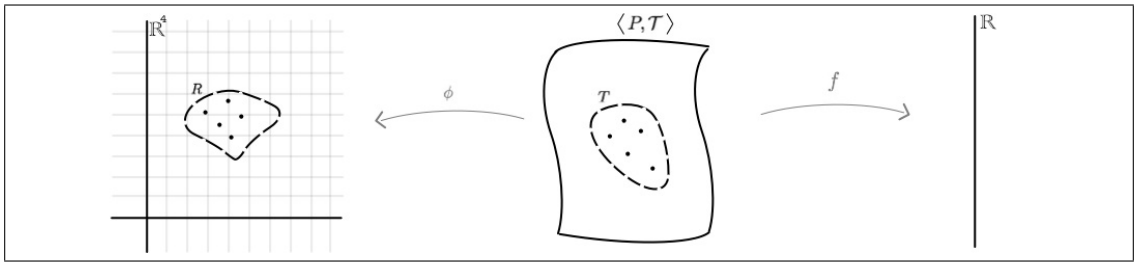


Figure 4: Pooley's material entity P, with a primitive Euclidean topological structure $\mathcal{T}$ (with open sets T). Its point-like parts have real-valued degrees of freedom, represented by f. A homeomorphism ϕ is shown mapping T onto R, the latter being an open subset of $\mathbb{R}^4$ with respect to the standard topology $\mathcal{R}^4$, which is introduced further in the discussion surrounding Figure 6(a).

relations captured in Leibnizian spacetime, Pooley notes that there is good reason to think that the same set of spatiotemporal relations would not support a best-systems account of the metrical structure of Minkowski spacetime as they do for the affine structure of Galilean spacetime. Instead, Pooley proposes taking a topological structure as primitive, and provides a simple illustration of how such a project would go.

Pooley's Illustration

Pooley describes a single material entity with point-like parts, the degrees of freedom of which can be modeled by a function into the reals—i.e. a single scalar field. He assumes the entity to be extended in four dimensions, and to have a structure that can be respected by some global mapping into $\mathbb{R}^4$—i.e. to have Euclidean topological structure.[12] As in Figure 4, his example can be illustrated by a topological structure $\mathcal{T}$ on a set P of field points, which together constitute a topological space $\langle P, \mathcal{T} \rangle$. The topological structure and the field points' degrees of freedom can be illustrated by a homeomorphism $\phi : \langle P, \mathcal{T} \rangle \to \mathbb{R}^4$ and a real-valued function $f : \langle P, \mathcal{T} \rangle \to \mathbb{R}$, respectively.[13]

With that Humean mosaic in mind, Pooley's liberalized best-systems approach to dynamical laws is similar to Huggett's, and can be summarized as follows. First, note that mappings like ϕ provide coordinatizations according to which the field points' degrees of

[12] Unless otherwise stated, $\mathbb{R}^4$ is assumed to have the standard Euclidean topology, $\mathcal{R}^4$, which is illustrated in Figure 6(a). In personal correspondence, Pooley has confirmed that with regard to Figure 4, he means for ϕ to be a homeomorphism into $\mathbb{R}^4$, making $\mathcal{T}$ a standard Euclidean topology.

[13] To allow for the possibility of holes in $\langle P, \mathcal{T} \rangle$, in which case a homeomorphism $\phi : \langle P, \mathcal{T} \rangle \to \mathbb{R}^4$ would not be possible, one might use the homeomorphism $\phi : \langle P, \mathcal{T} \rangle \to R$, where R is some open subset of $\mathbb{R}^4$.

freedom might be given a mathematical description. For instance, the values could be described by the function $f_\phi = f \circ \phi^{-1}$, so that different coordinatizations ϕ bring about different descriptions f_ϕ. Now, suppose that $\langle P, \mathcal{T} \rangle$ is such that given all homeomorphisms ϕ, there exists a collection $\mathbf{S}$ of coordinatizations with respect to which each corresponding function f_ϕ solves some kind of equation. Then, suppose further that for some collection $\mathbf{B} \subset \mathbf{S}$, the f_ϕ all solve some particular equation L, which has a better combination of simplicity and strength than the equations solved by the f_ϕ corresponding to the other coordinatizations in $\mathbf{S}$. If L is Lorentz-covariant, meaning that it preserves its form under elements of the Poincaré group, then Minkowski geometry could be said to supervene upon L's being the 'best' systematization of the field points' values and topological structure.[14] Put differently, if the coordinatizations $\phi \in \mathbf{B}$, according to which the f_ϕ solve L, are related to one another by elements of the Poincaré group, then Minkowski geometry would be "grounded in the existence of a proper subset of the coordinate systems adapted to the ~~Leibnizian relations~~ [topological structure] with respect to which the description of the entire ~~relational history~~ [topological space] is the solution of particularly simple equations" [Pooley 2013, pp. 571–2, my strikeouts].

Pooley is right that this construal of the dynamical approach would support the explanatory claims outlined in §1.3 above. It would also meet the criteria outlined in §2.3: neither Minkowski spacetime nor its point-like parts feature in the ontology; the spatiotemporal components of the ideology do not constitute Minkowski geometry; and Minkowski geometry codifies the behavior of material objects as described by the (liberalized best-system) Lorentz-covariant dynamical laws. However, there is no doubt that this proposal is somewhat bolder than Huggett's. That is, in accounting for Minkowski geometry, Pooley's primitives do not provide as strong a platform as do Huggett's in accounting for Galilean spacetime. For whereas Huggett proposes an account of Euclidean affine structure by taking as primitive material objects with mass and charge and all the metrical relations encoded by Leibnizian spacetime, Pooley proposes an account of

[14] Lorentz covariance and the Poincaré group are introduced toward the end §1.1, above.

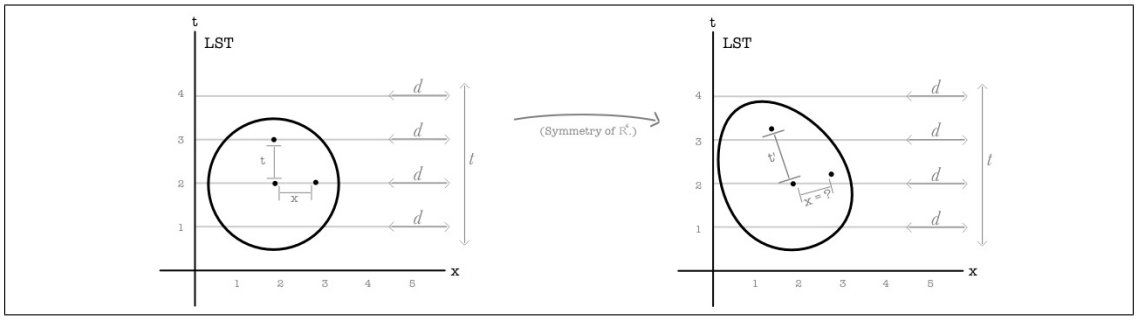


Figure 5: Euclidean topological structure ($\mathcal{R}^4$) is preserved by deformations that do not 'cut' or 'paste' the space, and under which Leibnizian relations may not be preserved. See also Figure 6(a).

the full metrical structure of Minkowski spacetime by taking as primitive a single set of scalar values endowed with mere Euclidean topological structure, which is far more symmetric than Leibnizian spacetime (see Figure 5).

Of course, Pooley's example is provided merely for the purposes of illustration, and he admits that a more realistic model would not be anything so simple. But in the search for a supportive example of Pooley's illustration, one may as well keep things as simple as possible. With that goal in mind, one thing is clear: whatever scalar values are taken as primitive, Pooley's best-systems claim would be on surer footing if some richer geometrical structure were taken as primitive. In particular, while Pooley opts for a primitive Euclidean topological structure, there might also be other, richer topologies worth considering, which would narrow the gap between what structure is taken as primitive, and what is to be explained away by an appeal to the symmetries of liberalized, best-system laws. Some examples are considered below.

The Induced and Alexandrov Topologies

A natural topology to consider for the constructivist project would be the 'induced' topology of Minkowski spacetime, the definition of which requires the introduction of one new term. Let an 'open sphere' S of a metric space $\langle P, d\rangle$ be the set of points that lie less than some given distance ϵ from a specified point p.[15] That is, $S(p,\epsilon) = \{q \mid q \in P,\ d(q,p) < \epsilon\}$. Then, the topology 'induced' by a metric space is the topology whose

[15]This description of induced topologies draws from Lipschutz's introduction to topology [1965, Ch. 8].

(a) In $\mathbb{R}^4$, the Euclidean metric d can be used to define 'open spheres' S, which generate the open sets R of the standard Euclidean topology $\mathcal{R}^4$. See also footnote 12 on page 107, and the discussion in Appendix A.

(b) The 'open hyperbolae' of the Minkowski 'metric' do not satisfy the axioms of topology.

Figure 6: The topologies induced by (a) the Euclidean metric d and (b) the Minkowski metric η.

open sets are generated by the space's open spheres. This means that the open sets of the induced topology are all and only those subsets which are unions of open spheres.[16]

Since a topology is sought such that, when it is assigned to a set of scalar field points, the field points' values are best systematized in terms of a Lorentz-covariant equation, a natural approach would be to explore the topology induced by the Minkowski metric, and then compare it to the Euclidean topology assumed by Pooley. But unfortunately, as illustrated in Figure 6(b), the open spheres of the Minkowski metric do not generate a topology; the sets they generate do not satisfy the basic axioms of a topological space.[17] This is a result of the fact that the Minkowski 'metric' η is not, strictly speaking, a metric at all. Indeed, it is not even a pseudo-metric: it not only allows distinct points to be separated by a 'distance' of zero, but it also fails to satisfy the triangle inequality.[18] It is

[16] That is, the open spheres form a 'base' for the induced topology [Lipschutz 1965, p. 86, p. 114]. Even the empty set, which, by the third axiom of topology (Appx. A), is always open, can be construed as a union of open spheres in the following way: $\emptyset$ is the union of the empty set of open spheres.
[17] The axioms of point-set topology are introduced in Appendix A.
[18] For a discussion, see the related work by Winnie [1977, p. 155].

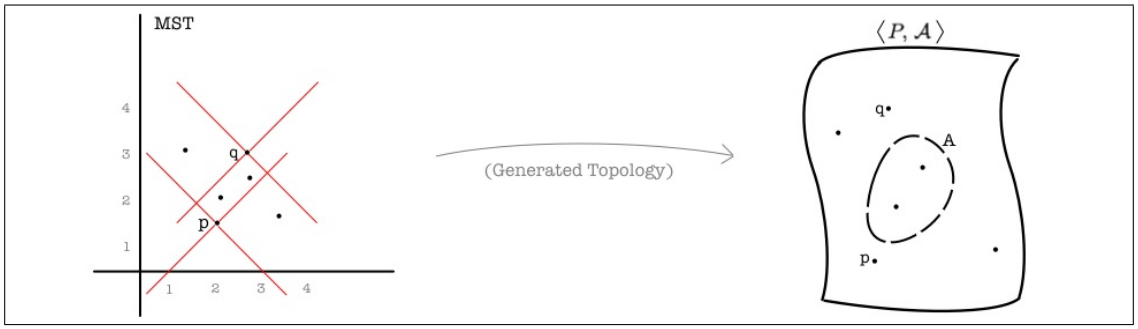


Figure 7: The Alexandrov 'interval' topology $\mathcal{A}$ is generated by the intersections of the chronological future of p and the chronological past of q, for all pairs of points (p, q) for which p is causally prior to q. (When q is on the light cone of p, the intersection is $\emptyset$, which is open in any topology.)

therefore standard practice to refer to the topology 'induced' by Minkowski spacetime as the topology induced by the Euclidean metric defined by the coordinates of any inertial frame. Naturally, this turns out to be $\mathcal{R}^4$, but the process doesn't seem to do any justice to the salient features of Minkowski geometry.

Now, while an induced topology is not strictly possible using the Minkowski metric, there are other ways to define the open sets of a topology that would seem to reflect at least some characteristics of Minkowski geometry, potentially making Pooley's proposal a bit less bold. For example, the Alexandrov topology $\mathcal{A}$ is illustrated in Figure 7, and has open sets A that are generated by the interiors of the 'Alexandrov intervals' of all pairs of causally connectible points. Given the centrality of causal structure to Minkowski spacetime structure, one might expect this topology to encode more of Minkowski geometry than the standard Euclidean topology.[19] However, the causal structure of Minkowski spacetime is actually such that the open sets of the Alexandrov topology generate the same topology as the open spheres of the Euclidean metric, and thus the Alexandrov topology for Minkowski spacetime coincides with $\mathcal{R}^4$.[20] While this suggests that $\mathcal{R}^4$ is more closely related to Minkowski geometry than it might at first seem, this has not stopped certain mathematicians from developing other, richer topologies for Minkowski spacetime. A noteworthy example is the 'fine' topology $\mathcal{F}$, introduced by Erik

[19] Regarding the centrality of causal structure to Minkowski metrical structure, earlier work by Robb [1914, 1936] and Zeeman [1964] is brought together nicely by Winnie [1977]. Giulini also offers a more recent discussion [2010, §4.2.4].

[20] This is discussed further by Winnie [1977, §IV], and also by Hawking and Ellis [1973, pp. 196–7].

Zeeman.

The Fine Topology

While $\mathcal{R}^4$ is a simple and familiar sub-metrical structure with which to work, it is too symmetric a space to capture any of the interesting structural features of Minkowski spacetime. For reasons unrelated to the project at hand, Zeeman sought a topology which could serve as a more fitting sub-metrical structure for Minkowski spacetime by having as its symmetry group something closer to that of Minkowski spacetime [1967]. Thus Zeeman introduced what he called the Fine topology $\mathcal{F}$, which can be defined by a relatively straightforward comparison with Euclidean topology. The standard Euclidean topology of Minkowski spacetime induces $\mathcal{R}^1$ on all of its time axes, and $\mathcal{R}^3$ on all of its space axes. $\mathcal{F}$ is simply defined as the finest topology to do the same. (Hence the name.) Thus $\mathcal{F}$ is finer than $\mathcal{R}^4$, meaning that $\mathcal{F}$ has all of the open sets of $\mathcal{R}^4$, and then some more—i.e. $\mathcal{R}^4 \subset \mathcal{F}$.

Why Zeeman took $\mathcal{F}$ to be a more suitable choice of topology for Minkowski spacetime becomes clear by exploring its open sets. The following process, illustrated in Figure 8, describes a collection of sets that are open in $\mathcal{F}$, but not in $\mathcal{R}^4$. First, consider Minkowski spacetime with the standard 'induced' Euclidean topology $\mathcal{R}^4$. Then, for some point p, choose an open set R containing p, and remove from R all other points lying on the light cone of p. Call the resulting subset F, so that $F = \{p\} \cup \{q \mid q \in R,\ I(p,q) \neq 0\}$, where $I(p,q)$ is the interval between p and q.[21] The same process can be repeated for any $R \ni p$, and then for any other point p in Minkowski spacetime. The resulting Fs will all be open in $\mathcal{F}$, but not in $\mathcal{R}^4$. Thus they are examples of the many sets that must be added to $\mathcal{R}^4$ in order to 'refine' it into $\mathcal{F}$.

It was one of Zeeman's primary results [1967, p. 161, p. 169] to show that the open sets of $\mathcal{F}$ enable one to deduce the light cone for any point in a topological space $\langle P, \mathcal{F} \rangle$. Then, since the causal structure of Minkowski spacetime is so closely related to its metrical

[21] Zeeman gives a similar introduction [1967, §5]. A thorough discussion is offered by Naber [2012, Appx. A].

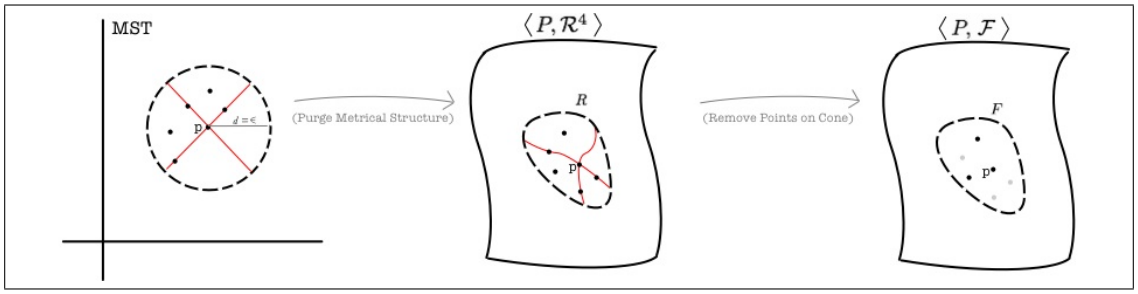


Figure 8: The Euclidean metric d on inertial coordinates of Minkowski spacetime induces $\mathcal{R}^4$. Removing from an open set R the points on the light cone of some point p (shown in red) creates a new set F, which is open in $\langle P, \mathcal{F} \rangle$, but not in $\langle P, \mathcal{R}^4 \rangle$.

structure, it should not be surprising that Zeeman shows the symmetry group of $\mathcal{F}$ to be the Poincaré group, plus dilations [§7]. This means that if a space $\langle P, \mathcal{F} \rangle$ is coordinatized by some mapping $\phi : P \to \mathbb{R}^4$, so that the open sets of $\mathcal{F}$ are mapped to open sets of $\mathbb{R}^4$, then the only coordinate transformations that would respect the space's open-set structure, so that open sets of $\mathcal{F}$ continue to be mapped to open sets of $\mathbb{R}^4$, will be those transformations that are elements of the Poincaré group and dilations.[22] But while this would give Lorentz-covariant equations a great advantage in systematizing any physical field with topological structure $\mathcal{F}$, this also means that the fine topology overshoots the goal for the 'constructivist' project at hand. As in the above quotation of Huggett, one's choice of topological structure must make it the case that "the topology of the fields plus point coincidences do not determine the metrical properties, so [that] there is additional work to be done by the laws" [2009, p. 418]. And while the Fine topology does not determine Minkowski interval relations between points in the space, the only degree to which the fine topology falls short of Minkowski geometry itself is with respect to scale. Thus the only work left for the laws to do, assuming they are not scale-invariant, would be to pick out a particular global scale. And so for a regularity relationalist account of Minkowski geometry, $\mathcal{F}$ is simply too rich a structure to take as primitive. But of course it was Zeeman's goal to find a topology with a symmetry group closer to that

[22] $\mathcal{F}$ being finer than $\mathcal{R}^4$ means that some such ϕ will not be bicontinuous, and thus not a homeomorphism. At best, ϕ will be bijective and continuous. (Indeed, any homeomorphism from $\langle P, \mathcal{R}^4 \rangle$ to $\mathbb{R}^4$ will suffer the same fate when $\mathcal{R}^4$ is 'refined' into $\mathcal{F}$ by adding open sets like the ones described above.) Or of course, $\mathbb{R}^4$ might be assigned some non-standard topology, so that ϕ could again be a homeomorphism.

of Minkowski spacetime, and so it should be unsurprising that it amounts to too rich a structure for the constructivist project.

While the three topologies outlined above do not amount to an extensive search for alternatives to $\mathcal{R}^4$ for the constructivist project, the results do not encourage a further search for any known topology that only partially closes the gap between $\mathcal{R}^4$ and Minkowski geometry. For whereas Galilean spacetime can be construed as a set of points on which one ascribes first a simultaneity relation, and then a temporal metric across those simultaneity slices, and then a spatial metric along the simultaneity slices, and finally a preferred affine connection and standard of rotation, Minkowski spacetime is constituted by a set of points to which one assigns a single spatiotemporal metric. Thus the relative simplicity of Minkowski metrical structure makes it difficult to improve upon $\mathcal{R}^4$ without arriving at a topology that actually constitutes Minkowski geometry itself. Of course, this is not to say that some such topology cannot be developed, and in fact it is suggested in §6.3 that there may be some low-hanging fruit in the field of developing a more fitting topology for a regularity relationalist construal of the dynamical approach. But in the meantime, $\mathcal{R}^4$ has not proven insufficient in any way; the only motivation in seeking an alternative was that the constructivist project might be made easier by choosing a richer topology. Lacking any such options for the time being, $\mathcal{R}^4$ is taken as the primitive geometrical structure in the following attempts to find a supportive example of a topologically structured scalar field that is arguably best-systematized by a Lorentz-covariant equation.

4.3 The Constructivist Project

Up to this point, it has been argued that topology would make a natural geometrical structure to assign to one's Humean mosaic in giving a best-systems account of Minkowski spacetime structure, and that despite its enormous symmetry group, $\mathcal{R}^4$ is a familiar and suitable topology with which to work. The 'constructivist' project can now be outlined in more detail. For simplicity, the following summary is based upon a particularly simple

system involving only the values of a single scalar field.[23] And, given that a set of scalar field values is sought which is best systematized by a Lorentz covariant equation, a natural place to begin the search for some such set would be among scalar-field solutions to Lorentz-covariant equations.

So, consider some solution ψ to a particular Lorentz-covariant equation E. For present purposes, it is useful to think of the solution ψ as describing a physical scalar field, understood in this context as a set P of field points having scalar-valued degrees of freedom and standing in Minkowski interval relations.[24] Thus let $\langle P, \eta \rangle$ be the space of scalar field points, endowed with Minkowski interval relations I, that constitute the physical field described by ψ. As illustrated in Figure 9, the field points' degrees of freedom can be represented by what will be called an 'adapted' real-valued function $s : \langle P, \eta \rangle \to \mathbb{R}$. The function s earns its name by the way it encodes the relative magnitudes of the field points' values—namely, so that for any two field points $p, q \in \langle P, \eta \rangle$, their image points $s[p], s[q] \in \mathbb{R}$ are such that the quotient $s[p]/s[q]$ is equal to the ratio of the values of p and q. Similarly, the field points' metrical structure can be represented by what will be called an 'adapted' coordinatization $\chi : \langle P, \eta \rangle \to \mathbb{R}^4$, which earns its name by the way it encodes the field points' metrical structure—namely, so that for any two field points $p, q \in \langle P, \eta \rangle$, their image points $\chi[p], \chi[q] \in \mathbb{R}^4$ are such that the multivariable scalar function $ds^2 = -(t_p - t_q)^2 + (x_p - x_q)^2 + (y_p - y_q)^2 + (z_p - z_q)^2$ returns a value equal to that of the interval relation I in which points p and q stand.[25] Then the field points can be given a coordinate-dependent mathematical description by taking the composition of these two mappings, $s \circ \chi^{-1} : \mathbb{R}^4 \to \mathbb{R}$. Indeed, this description is ψ, mentioned above.[26]

But what of the question at hand? The dynamical approach does not take as primitive some entity that could be modeled as $\langle P, \eta \rangle$. Rather, it was proposed that the constructivist

[23] In Appendix B, it is shown how the constructivist project can also be carried out with multiple fundamental physical fields. Instead of assuming the field-points to form material events by standing in some coincidence relation, it is argued that coincidence can be reduced to topological indiscernibility.

[24] This interpretation takes physical fields to be entities in their own right. This is discussed further in §6.1.

[25] As suggested in Figure 9, let the axes of $\mathbb{R}^4$ be labeled t, x, y, and z.

[26] Note the many-to-one relationship between ψ and $\langle P, \eta \rangle$. For any $\langle P, \eta \rangle$, there are many possible adapted coordinatizations χ and adapted functions s. Every unique pair (χ, s) results in a unique ψ.

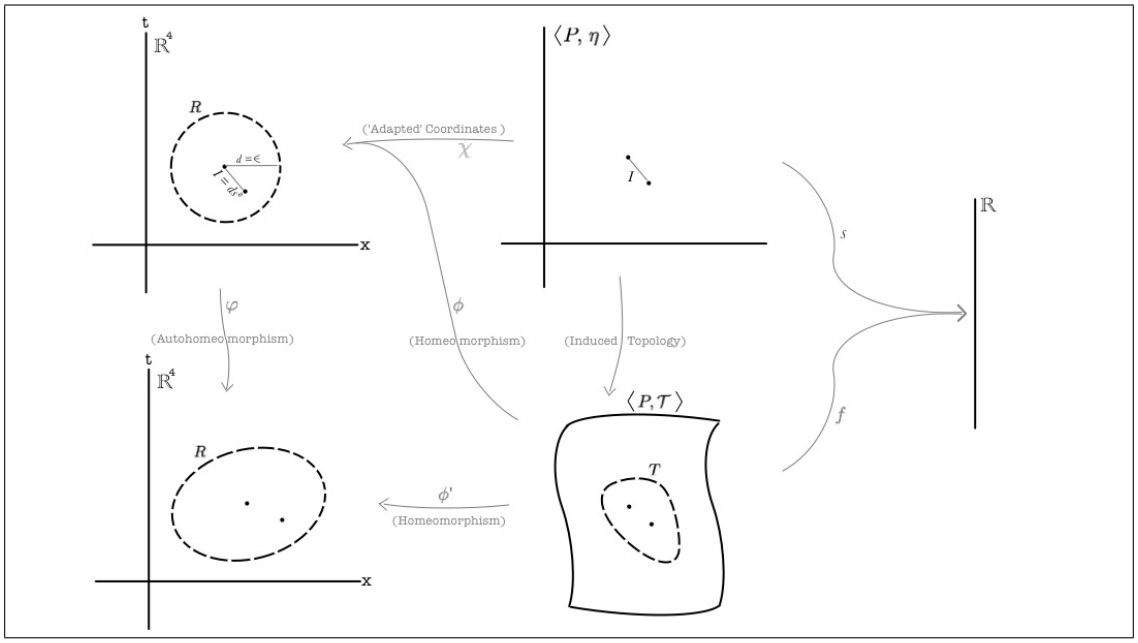


Figure 9: When a physical field $\langle P, \eta \rangle$ is coordinatized so that $I = ds^2$, the Euclidean metric d induces a Euclidean topology $\mathcal{T}$ with open sets T. The topological space $\langle P, \mathcal{T} \rangle$ is a candidate example for the constructivist project. (*cf.* Fig. 4.) It can be coordinatized by the homeomorphism $\phi : \langle P, \mathcal{T} \rangle \to \mathbb{R}^4$ as well as by any homeomorphism $\phi' = \varphi \circ \phi$, where $\varphi : \mathbb{R}^4 \to \mathbb{R}^4$ is an autohomeomorphism. Note that if $\psi = s \circ \chi^{-1}$ solves an equation E, so does $f_\phi = f \circ \phi$. That is, *qua* mappings from $\mathbb{R}^4$ to $\mathbb{R}$, $f_\phi = \psi$.

take as primitive some four-dimensional, topologically structured entity. So, recall from the discussion surrounding Figure 6(a) that the choice of a Euclidean metric on adapted coordinates χ does in some sense 'induce' a Euclidean topology for the field points in $\langle P, \eta \rangle$. Let this Euclidean topology be called $\mathcal{T}$. Thus it is possible to consider the field points with respect to their topological structure alone by referring to the topological space $\langle P, \mathcal{T} \rangle$. Indeed, this will be the starting point in choosing a candidate example for Pooley's illustration: assume the existence of the set P of field points described by a solution ψ to some Lorentz-covariant equation E, and then endow those field points with this topological structure $\mathcal{T}$. In what follows, $\langle P, \mathcal{T} \rangle$ is used as shorthand for the topological space described in the previous sentence. In these terms, the question at hand is whether there are any examples of a Lorentz-covariant equation E with solutions such that $\langle P, \mathcal{T} \rangle$ cannot be systematized by any other equations with a different covariance group than E, and with a greater or equal balance of simplicity and strength than E. A few more words about Figure 9 will clarify what is meant by systematization and

covariance in this context.

Recall from Figure 4 (p. 107) that for some such $\langle P, \mathcal{T} \rangle$, the field points' values can be represented by a mapping $f : \langle P, \mathcal{T} \rangle \to \mathbb{R}$. For simplicity, let f map each field point to the same location in $\mathbb{R}$ as does s, above. As for their topological structure $\mathcal{T}$, this can be represented by a homeomorphism into $\mathbb{R}^4$. For example, let the homeomorphism $\phi : \langle P, \mathcal{T} \rangle \to \mathbb{R}^4$ in Figure 9 map the field points to the same positions in $\mathbb{R}^4$ as the aforementioned adapted coordinatization χ.[27] Under this coordinatization, the values of the field points in $\langle P, \mathcal{T} \rangle$ can be given a coordinate-dependent description $f_\phi = f \circ \phi^{-1}$. Indeed, there will be a wide range of coordinate-dependent descriptions of the field points' values: for each adapted function f, there will be one mathematical description f_ϕ for every possible coordinatization of $\langle P, \mathcal{T} \rangle$. However, when it comes to systematizing the space, only a subset of those possible coordinatizations will be permissible—namely, those which 'respect', or 'preserve' the field points' Euclidean topology $\mathcal{T}$ by being homeomorphisms between $\langle P, \mathcal{T} \rangle$ and $\mathbb{R}^4$. Any other coordinatization would produce coordinate-dependent descriptions that fail to properly encode the primitive 'ordering' or 'arrangement' that is the field points' Euclidean topological structure. As in Figure 9, the set of permissible alternative coordinatizations for $\langle P, \mathcal{T} \rangle$ comprises the coordinatizations ϕ' that result from any autohomeomorphism $\varphi : \mathbb{R}^4 \to \mathbb{R}^4$, so that $\phi' = \varphi \circ \phi$.[28] Thus the set of permissible coordinate-dependent descriptions of the field points' values will comprise f_ϕ and every $f_{\phi'}$ associated with some such ϕ'.

Back to the question at hand. What is sought is a supportive example for the constructivist project: a space $\langle P, \mathcal{T} \rangle$ such that the best systematization of the field points' values and Euclidean topological structure is some Lorentz-covariant equation. Note, however, that it is not the coordinate-dependent descriptions $f_{\phi'}$ of the field points' values that serve as the systematization. Rather, it is the equations that the $f_{\phi'}$ solve. In Figure 9, it should be clear from the discussion above that for the aforementioned

[27] That ϕ so defined would be a homeomorphism follows from the fact that χ was the coordinatization used above in establishing the field points' 'induced' topology $\mathcal{T}$.

[28] Note that such coordinatizations are themselves homeomorphisms from $\langle P, \mathcal{T} \rangle$ into $\mathbb{R}^4$.

homeomorphism ϕ, the corresponding f_ϕ solves the same equation E with which the discussion began. Furthermore, the fact that E is a Lorentz-covariant equation ensures that for any ϕ' related to ϕ by an autohomeomorphism φ in the Poincaré group, the resulting $f_{\phi'}$ will also solve E.[29] However, nothing has yet been said about the equations (if any) that are solved by the $f_{\phi'}$ associated with all other permissible alternative coordinatizations ϕ'. Are they too Lorentz-covariant? If not, how does their balance of simplicity and strength compare to that of E?

Such an evaluation would go as follows. For any permissible alternative coordinatization ϕ', it must be asked whether the corresponding $f_{\phi'}$ also solves any equations. If, for all ϕ', the answer is always either "No, $f_{\phi'}$ does not solve any equations", or else "Yes, $f_{\phi'}$ solves E and no other equations", then $\langle P, \mathcal{T} \rangle$ is systematized best—or in this case, *only*—by some Lorentz-covariant equation, E. In that case, every ϕ' with respect to which $f_{\phi'}$ solves E would be related to ϕ by some element of the Poincaré group, and $\langle P, \mathcal{T} \rangle$ would vindicate Pooley's illustration above.[30] However, whenever the answer is "Yes, $f_{\phi'}$ solves some other equation F", then F will have to be included in the running for the best-systems dynamical law of $\langle P, \mathcal{T} \rangle$. If E is deemed the winner, then once again $\langle P, \mathcal{T} \rangle$ is systematized best—though not *only*—by some particular Lorentz-covariant equation. Some such $\langle P, \mathcal{T} \rangle$ would vindicate Pooley's illustration equally well. But finally, in the case that some other equation F wins the title, one should then inquire after the symmetry group of F, which would be determined by the transformations under which F preserves its form. If that symmetry group is the Poincaré group, then yet again $\langle P, \mathcal{T} \rangle$ is systematized best by some Lorentz-covariant equation. However, if the symmetry group of this rival equation F differs from that of E, then $\langle P, \mathcal{T} \rangle$ does not turn out to be supportive of the constructivist project; the metrical structure that supervenes

[29] That is, such an $f_{\phi'}$ will be another ψ. See footnote 26 on page 115, and also the caption of Figure 9.

[30] This claim requires a qualification. Certainly, if E is a Lorentz-covariant equation solved by f_ϕ, then whenever φ is an element of the Poincaré group, the corresponding $f_{\phi'}$ will also solve E. After all, part of what it is for the Poincaré group to be the symmetry group of Lorentz-covariant dynamical laws is for elements of that group to map solutions to solutions and non-solutions to non-solutions. However, this does not rule out the possibility that some transformation φ might *not* be a member of the Poincaré group and yet still map a solution of E to another solution. In what follows, such fluke transformations are ignored.

upon the symmetries of the best-system dynamical law F would be something other than Minkowski geometry.

In the sections that follow, solutions to familiar Lorentz-covariant field equations are considered as candidate examples of such a project. But before turning to those examples, one point should be stressed. An advocate of the dynamical approach should not be surprised or concerned if supportive examples are difficult to find in this context. A single scalar field with Euclidean topological structure may be too simple and symmetric to support the kind of best-systems claim under consideration. Indeed, this was the concern that inspired the discussion in §4.2. Should that be the case, the going construal of the dynamical approach is certainly not doomed, for it may still be the case that Minkowski geometrical structure could be shown to match the symmetries of the best-systems dynamical laws for a topologically structured *set* of scalar fields, or for a topologically structured set of vector fields or tensor fields. But for now, all discussion of multiple scalar fields, or of any number of vector or tensor fields, is postponed; the following chapter serves as a simple case-study for the claim at the heart of this construal of the dynamical approach.

5 A Case Study in Regularity Relationalism: Scalar Fields with Euclidean Topological Structure

In the preceding chapter, the dynamical approach to relativity was described as a relativistic version of Nick Huggett's 'regularity relationalism', which would give a 'liberalized' best-systems account of Minkowski spacetime structure. The proposed geometrical structure of the constructivist's Humean mosaic was the standard Euclidean topology of $\mathbb{R}^4$. In what follows, this chapter explores just how difficult it would be to come up with examples of physical fields, the values of which, when endowed with the standard Euclidean topological structure, vary in a sufficiently non-symmetric way so as to be arguably best-systematized by some Lorentz-covariant equation.

The question at hand, as outlined in §4.3, is this: are there any examples of a four-dimensional scalar-field solution to a Lorentz-covariant equation, such that the solution could not, under an appropriate homeomorphism, solve some otherwise-covariant equation with a greater or equal balance of simplicity and strength? Such a scalar field, when considered up to homeomorphism as $\langle P, \mathcal{T} \rangle$, would serve as a supportive example for the so-called 'constructivist project' of the dynamical approach. Unfortunately, a definitive 'Yes' or 'No' answer to this question is beyond the scope of this thesis. The standard Euclidean topology $\mathcal{R}^4$ has a sufficiently large symmetry group to prohibit any attempt to outline all possible $f_{\phi'}$ for some given scalar field, let alone what equations the $f_{\phi'}$ might solve. And even if such a project could be carried out, there would surely arise situations that call for more robust definitions of 'simplicity' and 'strength' than is provided here. Still, some insight can be gained by asking the question with regard to the basic solutions of a few familiar Lorentz-covariant equations. And in doing so, it is shown that although the simplest solutions to these familiar equations clearly fail as supportive examples of the constructivist project, one does not have to venture far into the equations' more complicated solutions to find examples in which the constructivist's best-systems claim begins to appear much more plausible—so plausible, in fact, that the onus is on the critic of the dynamical approach to come up with a rival systematization of these solutions,

upon which Minkowski geometrical structure would not supervene.

To begin, some familiar scalar field equations are introduced in §5.1. That the simplest solutions to those equations do not provide supportive examples for the constructivist project is established in §5.2. It is then shown in §5.3 that only slightly more complicated solutions are necessary in order for the constructivist's project to look much more promising. Finally, the lessons learned are summarized in §5.4.

5.1 The Klein–Gordon Equation and Its Solutions

As explained in the preceding chapter, a candidate example for the constructivist project of the dynamical approach is an entity $\langle P, \mathcal{T} \rangle$ that might be described as a Euclidean topological space of real-valued scalar field points. A natural place to begin looking for some such $\langle P, \mathcal{T} \rangle$ is in the space of solutions to familiar Lorentz-covariant scalar field equations, such as the Klein–Gordon equation. Whether or not its solutions, when considered up to homeomorphism as $\langle P, \mathcal{T} \rangle$, admit permissible recoordinatizations ϕ' for which $f_{\phi'}$ solves any equations of a different covariance group, is addressed in the following sections. In preparation for that discussion, this section gives a brief derivation of the Klein–Gordon equation, and also of its non-relativistic cousin, the Schrödinger equation.[1] But first, it should be noted at the outset that in this section, all talk of quantum states is purely heuristic. The Klein–Gordon equation is of interest to us only as a Lorentz-covariant equation that is solved by a scalar function. And for reasons that become clear below, it will be useful to have compared its derivation to that of the Schrödinger equation.

To begin, recall the nonrelativistic energy–momentum relation for a free particle: $E = \frac{1}{2}mv^2 = \frac{1}{2m}p^2$. Recall also that energy and momentum can be turned into operators for a quantum state $\psi(t, \bar{x})$ via the standard correspondence relations $E \to i\partial_t$ and $p \to -i\nabla$.[2] Doing so gives us the time-dependent Schrödinger equation for a single, free,

[1] What follows is not a historical outline of the discovery of these equations. Armin Wachter provides a similar introduction, from which this section draws [2011, §1.1]. Alternatively, a derivation via the Lagrangian and Hamiltonian formulation for fields is given by Goldstein *et al.* [2002, §13.6].

[2] $\nabla = \partial_x + \partial_y + \partial_z$; $\nabla^2 = \nabla \cdot \nabla$. Note also that natural units are chosen, so that $c = \hbar = 1$.

nonrelativistic particle.

$$i\partial_t\psi + \frac{1}{2m}\nabla^2\psi = 0 \tag{3}$$

As its derivation began with a nonrelativistic energy–momentum relation, it is unsurprising that the the Schrödinger equation is not Lorentz-covariant.[3] But consider now the square of the relativistic energy–momentum relation, $E^2 = p^2 + m_0^2$.[4] The same correspondence relations and some rearranging of terms gives the Klein–Gordon equation for a single, free, relativistic particle.

$$-\partial_t^2\psi + \nabla^2\psi = m_0^2\psi \tag{4}$$

Note that the Klein–Gordon equation is indeed Lorentz-covariant. As explained in §1.1 (p. 9), this means that the Klein–Gordon equation—but not the Schrödinger equation—retains its form under any coordinate transformation that is an element of the Poincaré group.

Now, despite their different covariance groups, these two equations both admit solutions of the following form,

$$\psi(t, \bar{x}) = Ae^{i(\bar{p}\cdot\bar{x} - Et)}$$

where A is some scalar value. Solutions of this form describe a wave-like distribution of both real and imaginary scalar values, as illustrated in Figure 10. To make the wave-like characteristics of these solutions more explicit, and in order to tease out the slight differences between the solutions to the two equations above, this shared form of solutions can be re-expressed by using the Planck-Einstein and De Broglie relations $E = \omega$ and

[3] While there is no doubt that the Schrödinger equation is not Lorentz-covariant, its precise symmetry group is less straightforward. The Schrödinger equation is sometimes called 'Schrödinger-covariant', where the Schrödinger group is an adaption of the Galilean group, and resolves the issue of phase change under a Galilean coordinate transformation. See the discussions by Lévy-Leblond [1963, §III; 1967, §II]. More recently, Brown and Holland have shown that the Schrödinger equation is Galilean-covariant in a wider range of potentials than typically thought [1999].

[4] Starting with $E = \sqrt{p^2 + m_0^2}$ turns out to be problematic. See the references in footnote 1, above.

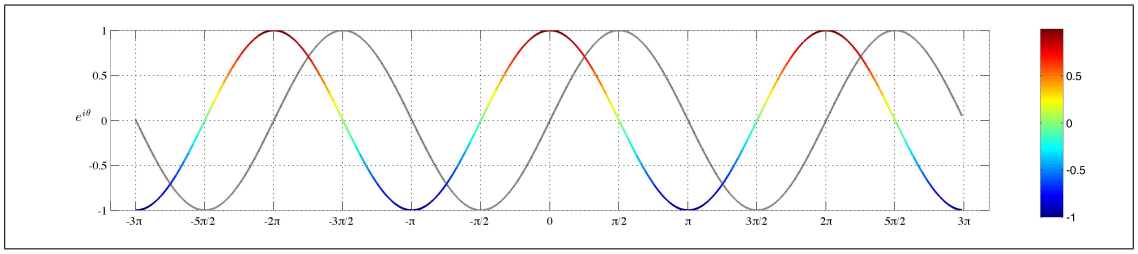


Figure 10: Real (in color) and imaginary (in grey) components of $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. As in all illustrations to follow, unless otherwise noted, the vertical axis represents the magnitude of the scalar function. For real-valued scalar functions, this magnitude corresponds to the color of the curve.

$\bar{p} = \bar{\kappa}$.[5]

$$\psi = Ae^{i(\bar{\kappa}\cdot\bar{x}-\omega t)} \tag{5}$$

Unless otherwise noted, ψ is used in the following sections to represent a solution of this form—a complex-valued 'planewave' solution. Planewaves have constant amplitude A, constant angular frequency ω, constant wave-vector $\bar{\kappa}$, and thus constant phase velocity $\bar{v}_p$. As illustrated in Figure 11, they also have a constant magnitude in all directions orthogonal to the wavevector.

Now, because the Planck-Einstein and De Broglie relations ensure that angular frequency ω and wavenumber κ are related in accord with the energy–momentum relation, planewave solutions to the Schrödinger and Klein–Gordon equations will differ precisely insofar as do the equations' so-called 'dispersion relations' between ω and κ. For the Schrödinger equation, applying the Planck-Einstein and De Broglie relations to the nonrelativistic energy–momentum relation gives one dispersion relation,

$$\omega = \frac{\kappa^2}{2m} \tag{6}$$

whereas applying the Planck-Einstein and De Broglie relations to the relativistic energy–momentum relation gives another for the Klein–Gordon equation.

$$\omega = \sqrt{\kappa^2 + m_0^2} \tag{7}$$

[5] As above, natural units are being used. Recall that $\omega := 2\pi\nu$ is the angular frequency of a wave, and that $\bar{\kappa}$ is the wave-vector, which points in the direction of the 'phase velocity' $\bar{v}_p$, the velocity of propagation of a single wave crest. The angular wave*number* is $\kappa = |\bar{\kappa}| := \omega/v_p$.

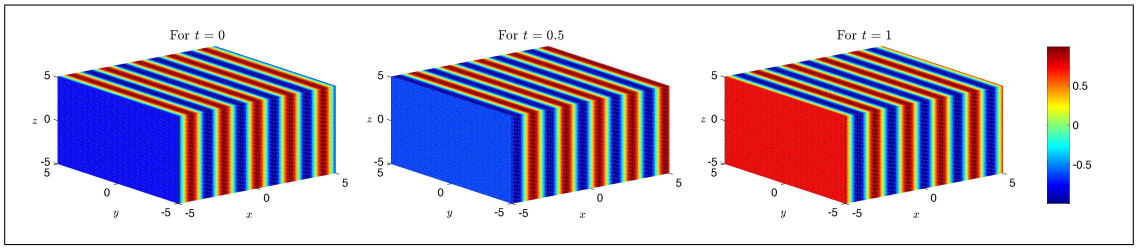


Figure 11: Three-dimensional snapshots of $\Re(\psi)$, the real-valued component of a planewave solution ψ, with wave-vector $\bar{\kappa}$ pointing along the x-axis (see footnote 5). For simplicity, $A = m = \hbar = 1$ and $\kappa = 3$. As illustrated, a slice of constant value (color) moves along the x-axis over time. Note that, unlike in other illustrations of this section, the vertical axis represents a third spatial dimension.

The differences in wave behaviour brought about by differences in dispersion relations will become clear below.[6]

In closing, it should be noted that while imaginary numbers may appear in ψ, they do not appear explicitly in the Klein–Gordon equation itself (Eqn. 4), whereas they do in the Schrödinger equation (Eqn. 3). Thus while the Schrödinger equation admits planewave solutions only in the form of Equation 5, the Klein–Gordon equation also admits planewave solutions that are entirely real-valued. Such solutions are still planewaves, in that they share the characteristics listed above. Indeed, the real-valued planewave solutions to the Klein–Gordon equation are nothing more than the real-valued components of the complex-valued planewave solutions ψ. Thus they can be written as such.

$$\psi_{\Re} = \Re\left(Ae^{i(\bar{\kappa}\cdot\bar{x}-\omega t)}\right) = A\cos(\bar{\kappa}\cdot\bar{x}-\omega t) \tag{8}$$

5.2 Ways to Systematize Planewaves

The Klein–Gordon equation is of interest as a Lorentz-covariant equation that is solved by a real-valued scalar function. Might its solutions serve as a supportive example of the constructivist project? That is, might any of its solutions be such that they could not, under some homeomorphism of $\mathcal{R}^4$, solve some otherwise-covariant equation with a greater or equal balance of simplicity and strength?

As a first step in putting into practice the ideas of §4.3, this section begins by

[6] In Equation 7, ω could equal the positive or negative of the root. Taking the negative is typically interpreted as representing an antiparticle, as discussed by Wachter [2011, §1.1].

addressing the question at hand with regard to the Klein–Gordon equation's complex-valued planewave solutions, of the form of Equation 5. This serves as a helpful exercise since in this context there is a 'rival' best-systems dynamical law that is very familiar—namely, the Schrödinger equation. Following that demonstration, the real-valued solutions to the Klein–Gordon equation are explored as potential supportive examples of the dynamical approach. Although the rival systematizations in that case are equally simple, they are less familiar equations, and each requires a brief introduction.

Naturally, one should expect that these simple, planewave solutions to the Klein–Gordon equation would not live up to the challenge. That is, it should be unsurprising to find that planewave solutions to the Klein–Gordon equation do indeed admit permissible coordinatizations ϕ' with respect to which $f_{\phi'}$ solves equations with different covariance groups. And as explained at the end of §4.3, this is of no concern for the constructivist. However, after showing how the simplest solutions to the Klein–Gordon equation fail to serve as supportive examples of the constructivist project, it is shown in §5.3 that these solutions need only become slightly more complicated in order to make the constructivist's best-systems claim highly plausible.

Complex-Valued Planewave Solutions

Let $P_{\mathbb{C}}$ be the set of field points that constitute the physical field represented by a complex-valued planewave solution ψ to the Klein–Gordon equation, as in Equation 5 (p. 124). When those field points are endowed with the 'induced' Euclidean topological structure $\mathcal{T}$, as described in §4.3, then there exists a homeomorphism $\phi : \langle P_{\mathbb{C}}, \mathcal{T} \rangle \to \mathbb{R}^4$, according to which f_ϕ serves as the real-valued component of a planewave solution to the Klein–Gordon equation.[7] And as explained at the end of §4.3, the first step in answering the question at hand will be to determine whether the Klein–Gordon equation is the only way to systematize $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$. If not, then there are 'rival' best-systems dynamical laws to

[7] Note that Figure 9 falls short in this context, since a complex-valued solution to the Klein–Gordon equation would have to be modelled by either *two* mappings, $\psi_{\Re} : \mathbb{R}^4 \to \mathbb{R}$ and $\psi_{\Im} : \mathbb{R}^4 \to \mathbb{R}$, or else by one mapping, $\psi_{\mathbb{C}} : \mathbb{R}^4 \to \mathbb{C}$.

be evaluated.

In fact, there are rival systematizations at play; the Klein–Gordon equation is not the only equation solved by the $f_{\phi'}$ corresponding to permissible coordinatizations ϕ' of $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$. For instance, there is a certain set of permissible coordinatizations for which the corresponding $f_{\phi'}$ solve the Schrödinger equation. To see this, recall from above that both the Klein–Gordon and Schrödinger equations admit planewave solutions ψ of the form of Equation 5, and that for any given choice of amplitude A, mass m, and wavenumber κ, the planewave solutions to these two equations will differ only insofar as do their dispersion relations. And what kind of change in wave behaviour is effected by a difference in dispersion relation? Obviously, wave equations with different dispersion relations will have different angular frequencies ω for some given wavenumber κ. But another difference follows: this change in angular frequency ω will not be accompanied by a corresponding change in wavelength λ, because $\kappa := 2\pi/\lambda$, and no change in wavenumber κ is being proposed. Rather, the difference in dispersion relation affects a change in angular frequency that is accompanied by a change of phase velocity $\bar{v}_p$, since $v_p = \omega/\kappa$.[8] Both of these differences are illustrated in Figure 12, which shows the real-valued components of planewave solutions to the Klein–Gordon and Schrödinger equations with equal-valued wavenumbers κ. The difference in angular frequency ω can be seen by the relative difference in the number of crests that fit within some time, and the difference in phase velocity $\bar{v}_p$ can be seen by the relative difference in the 'slope' of the wavecrests.

The important point is that differences of this sort can be accommodated by homeomorphisms of the field points' Euclidean topology. That is, in the language of §4.3, these differences can be brought about by an autohomeomorphism coordinate transformation φ. Indeed, both can be accounted for by one and the same uniform 'stretching' of the time coordinate. In particular, if the time coordinate of Figure 12-ii were stretched

[8] As in footnote 5, $v_p = \nu\lambda = 2\pi\nu/\kappa = \omega/\kappa$. When the dispersion relation is linear, a change in κ does not result in a change in v_p, but merely in λ. Having non-linear dispersion relations, the Klein–Gordon and Schrödinger equations are examples of so-called 'dispersive' wave equations.

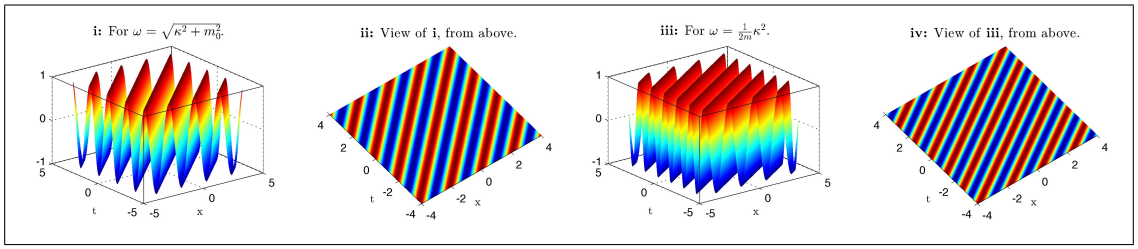


Figure 12: $\Re(\psi)$, the real-valued component of a planewave ψ. As in Figure 11, $A = m = \hbar = 1$ and $\kappa = 3$. Plots i and ii show $\Re(\psi)$ for the dispersion relation of the Klein–Gordon equation, while iii and iv show $\Re(\psi)$ for that of the Schrödinger equation. Plots ii and iv reveal a difference in $\bar{v}_p$, but not in λ.

by a small amount, the result would be the plot of Figure 12-iv.[9] The upshot of all this is that for any such $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$, which by definition admits a coordinatization ϕ with respect to which f_ϕ serves as the real-valued component of a planewave solution to the Klein–Gordon equation, there is a permissible alternative coordinatization ϕ' with respect to which $f_{\phi'}$ serves as the real-valued component of a solution to the Schrödinger equation. What's more, the same can be said about the imaginary components $\Im(\psi)$ of these complex-valued solutions. As suggested by Figure 10, and as illustrated in Figure 13, the magnitude of the imaginary component of a complex-valued planewave is out of phase with, but otherwise equal to, that of the real-valued component. Therefore the imaginary-valued component of complex-valued solutions to these equations needn't be given separate attention when considering coordinate transformations. For if some topology-preserving coordinate transformation φ maps a coordinatization ϕ (with respect to which f_ϕ serves as a real-valued component of a complex-valued planewave solution to the Klein–Gordon equation) into an alternative coordinatization ϕ' (with respect to which $f_{\phi'}$ serves as the real-valued component of a complex-valued planewave solution to the Schrödinger equation), then the same would be true with respect to the corresponding imaginary-valued components of the planewave solutions.[10] But all of this is simply to say that the $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$ under consideration here is just as easily systematized by the

[9] Specifically, the time dilation is: $t \to t' = \frac{\sqrt{\kappa^2 + m_0^2}}{\kappa^2/2m} t$. See also the the caption to Figure 19.

[10] Note that the phase difference between the real- and imaginary-valued components would also be preserved in such a transformation. As illustrated in Figure 10, the real and imaginary components are out of phase by one quarter of a wavelength, and as in the discussion surrounding footnote 8 on page 127, no change in wavelength (wavenumber) is being proposed.

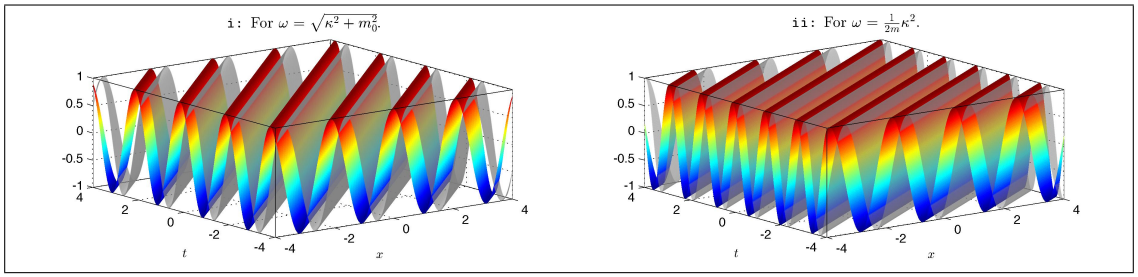


Figure 13: Two illustrations of a complex-valued planewave, ψ. In both, $\Re(\psi)$ is shown in color and $\Im(\psi)$ in grey. As above, $A = m = \hbar = 1$ and $\kappa = 3$. Plot i shows ψ for the dispersion relation of the Klein–Gordon equation. Plot ii shows ψ for the dispersion relation of the Schrödinger equation.

Schrödinger equation.

Following the programme laid out in §4.3, it now remains to be asked what the covariance group is for the Schrödinger equation. For present purposes, suffice it to say that it is not the Poincaré group; the Schrödinger equation is not Lorentz-covariant.[11] And because Klein–Gordon equation does not systematize $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$ in a way that is obviously simpler or in any way stronger than the Schrödinger equation does, this does not bode well for finding a single planewave solution to the Klein–Gordon equation that would serve as a supportive example of the constructivist project. In fact, many other equations could be given as rival systematizations of some such $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$, as shown below. But of course the Klein–Gordon equation admits not only complex-valued, but also real-valued planewave solutions. And although the systems under consideration are at this point extremely simple, it is worth showing in some detail that there are rival systematizations also for its real-valued planewave solutions.

Real-Valued Planewave Solutions

Consider now some real-valued planewave solution $\psi_{\Re}$ to the Klein–Gordon equation, as in Equation 8 (p. 125). As before, let P be the set of field points that constitute the physical field, and ascribe to those field points their 'induced' Euclidean topological structure $\mathcal{T}$. And once again, in order to determine whether the best systematization of $\langle P, \mathcal{T} \rangle$ is a Lorentz-covariant equation, it must first be determined whether there are any

[11] See footnote 3 on page 123.

permissible recoordinatizations ϕ' according to which $f_{\phi'}$ solves another equation.

To begin, it should be noted that because the Schrödinger equation does not admit real-valued planewave solutions, it is not in the running for the best-systems dynamical law of $\langle P, \mathcal{T} \rangle$. Still, this doesn't ensure the title for the Klein–Gordon equation. In fact there are other equations that admit real-valued solutions of this same form. And more importantly, $\langle P, \mathcal{T} \rangle$ does in fact admit permissible coordinatizations ϕ' with respect to which the $f_{\phi'}$ solve these other equations. In what follows, three examples are given.

The 'Squaringer' Equation: To introduce the first example, it will be useful to recall why it is that the Schrödinger equation (Eqn. 3, p. 123) admits solutions ψ of the form of Equation 5, but not real-valued solutions $\psi_{\Re}$ of the form of Equation 8. The first obstacle is apparent: the complex number i appears in the Schrödinger equation, and while it gets 'squared away' when ψ is a complex-valued planewave, this is not the case with $\psi_{\Re}$. This results in a complex-valued dispersion relation between ω and κ, which brings about amplifying and dampening effects on the solution, resulting in exponential behavior.[12] And so the first step in making the Schrödinger equation compatible with real-valued solutions would be to drop the i.

$$\partial_t \psi + \frac{1}{2m} \nabla^2 \psi = 0$$

Still more obstacles arise, since the first-order time derivative of $\psi_{\Re}$ results in a sine function, rather than the sought-after cosine function that is $\Re(\psi)$. This could be fixed by changing the first term to a second-order time derivative.

$$\partial_t^2 \psi + \frac{1}{2m} \nabla^2 \psi = 0$$

However, in order to avoid complex-valued dispersion relations once again, it would be necessary to change the sign of one of the two terms. The result would not be far from the Klein–Gordon equation itself.[13]

[12] Some examples are illustrated below. See, for instance, Figures 24 and 25.
[13] Indeed, aside from the constant $1/2m$, the result is the free wave equation, introduced below.

The above should be no surprise: the steps just rehearsed effectively make up for the difference in the energy–momentum relations from which the Schrödinger and Klein–Gordon equations were derived in §5.1. To see this, recall that whereas the Schrödinger equation was derived from the nonrelativistic energy–momentum relation $E = \frac{1}{2m}p^2$, the Klein–Gordon equation was derived from the *square* of the relativistic energy–momentum relation: $E^2 = p^2 + m_0^2$. And while despite this difference both equations admit solutions of the form of ψ (Eqn. 5), it is precisely this difference which enables the Klein–Gordon equation to admit solutions of the form of $\psi_{\Re}$ (Eqn. 8). Squaring the energy term ($E \to i\partial_t$) causes the i to be 'squared away', which gives the change of sign rehearsed above, and also provides the second-order time derivative necessary to recover a cosine term. Indeed, consider if the nonrelativistic energy–momentum relation, from which the Schrödinger equation was derived, were given the same 'squaring' treatment.

$$E = \frac{p^2}{2m} \quad \longrightarrow \quad E^2 = \frac{p^4}{4m^2}$$

Applying the same correspondence relations ($E \to i\partial_t$ and $p \to -i\nabla$) would result in what I will call the 'Squaringer' equation.

$$\partial_t^2\psi + \frac{1}{4m^2}\nabla^4\psi = 0 \tag{9}$$

Like the Klein–Gordon equation, this admits both complex- and real-valued planewave solutions.

Now, for both the real- and complex-valued solutions of the Squaringer equation, the dispersion relations match those of the Schrödinger equation: $\omega = \kappa^2/2m$. And above, it was shown that complex-valued planewaves with this dispersion relation are homeomorphic to solutions to the Klein–Gordon equation. The same is therefore true of real-valued planewaves with this dispersion relation. That is, when a single, real-valued planewave solution to the Klein–Gordon equation is considered up to homeomorphism as $\langle P, \mathcal{T}\rangle$, there are permissible alternative coordinatizations ϕ' with respect to which $f_{\phi'}$ will solve the Squaringer equation. And because the Squaringer equation is not Lorentz-

covariant, it serves as a threat to the claim that $\langle P, \mathcal{T} \rangle$ might be systematized best by a Lorentz-covariant equation. In fact, the derivation of the Squaringer equation paves the way for generating other rival systematizations involving partial differential equations of arbitrarily high order—for what if the nonrelativistic energy–momentum relation had been given, say, two 'squaring' treatments? The resulting equation would admit both real- and complex-valued planewave solutions, though the higher-order exponents might come at a cost in simplicity. Indeed, the same argument might also be made against the Squaringer equation itself, but of course the difference in order is not great.

The Free Wave Equation: By making small changes to the Klein–Gordon equation itself (Eqn. 4, p. 123), other rival systematizations of $\langle P, \mathcal{T} \rangle$ can be found. Consider for example the free wave equation, which is effectively the Klein–Gordon equation with $m_0 = 0$.

$$-\partial_t^2 \psi + \nabla^2 \psi = 0 \tag{10}$$

Just like the Klein–Gordon, Schrödinger, and Squaringer equations, the free wave equation admits complex-valued planewave solutions ψ of the form of Equation 5. As one might expect, it also admits real-valued solutions $\psi_{\Re}$ of the form of Equation 8. And once again, the $\langle P, \mathcal{T} \rangle$ chosen above admits permissible coordinatizations ϕ' according to which $f_{\phi'}$ solves the free wave equation.

To see this, note first that the free wave equation's real- and complex-valued planewave solutions have a dispersion relation of $\omega = \kappa$.[14] Then, simply recall the two lessons from above: (i) that given a planewave solution with some particular amplitude A, mass m_0, and wavenumber κ, a change in dispersion relation effects nothing more than a change in the wave's angular frequency ω and phase velocity $\bar{v}_p$; and (ii) that these differences can be accommodated by symmetries of the field points' Euclidean topological structure $\mathcal{T}$—i.e. by permissible coordinate transformations φ. Indeed, as illustrated in Figure 14, real-valued planewave solutions of the free wave equation are related to those of the

[14] More precisely, $\omega = \pm\kappa$, but see footnote 6 of this chapter. Note also that the free wave equation is not 'dispersive'; changing κ affects ω and λ, but not v_p. (See also footnote 8 on page 127.)

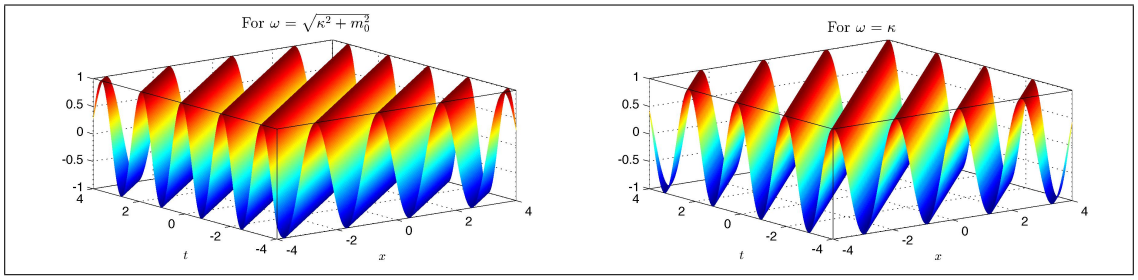


Figure 14: Two illustrations of $\Re(\psi)$ for a planewave ψ. On the left, like Figure 12-i, is $\Re(\psi)$ for the dispersion relation of the Klein–Gordon equation. On the right is $\Re(\psi)$ for that of the free wave equation. As in earlier figures, $A = \hbar = 1$ and $\kappa = 3$. However, in order to make the differences visible, $m_0 = 2$.

Klein–Gordon equation by the same kind of transformations as were the real-valued components of complex-valued solutions of the Klein–Gordon and Schrödinger equations in Figure 12.[15]

Now, despite the striking similarity between the free wave and Klein–Gordon equations, the free wave equation is not, strictly speaking, Lorentz-covariant. For while the free wave equation does indeed preserve its form under elements of the Poincaré group, its having no mass term also gives rise to additional symmetries.[16] Therefore if the free wave equation were deemed the simplest and strongest systematization of $\langle P, \mathcal{T} \rangle$, then its having a larger symmetry group than the Klein–Gordon equation means that the supervenient geometrical structure would be weaker than that of Minkowski spacetime.

The 'Schmein–Gordon' Equation: By considering other adaptations to the Klein–Gordon equation, another rival systematization can be found. Consider for example what I will call the 'Schmein–Gordon' equation, which results from changing the sign of certain terms in the relativistic energy–momentum relation from which the Klein–Gordon equation was derived, above.

$$E^2 = p^2 + m_0^2 \;\rightarrow\; E^2 = -p^2 + m_0^2 \tag{11}$$

[15] With respect to its complex-valued solutions, the same comparison could be made to Figure 13.

[16] The symmetries of the free wave equation (in two spatial dimensions) are rehearsed by Olver [1986, §2.4] and by Miller [1977, Ch. 4]. They include not only the translations and hyperbolic rotations of the Poincaré group, but also global dilations, 'inversions', and the 'special conformal transformations'.

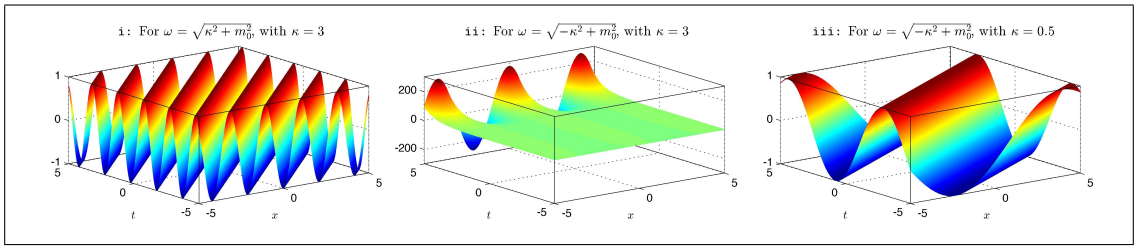


Figure 15: Three illustrations of $\Re(\psi)$ for a planewave ψ. As in earlier figures, $A = \hbar = m_0 = 1$. As in Figures 12, 13, and 14, Plot i shows a real-valued planewave solution to the Klein–Gordon equation, with wavenumber $\kappa = 3$. Plot ii shows the real-valued component of a complex-valued solution to the Schmein–Gordon equation with $\kappa^2 > m_0^2$. (The equation admits no real-valued solutions in that context.) Plot iii shows a real-valued planewave solution to the Schmein–Gordon equation with $\kappa^2 \leq m_0^2$.

The result differs from the Klein–Gordon equation only by a corresponding sign change.

$$\partial_t^2 \psi + \nabla^2 \psi = -m_0^2 \psi \tag{12}$$

Like the Klein–Gordon, Squaringer, and free wave equations, the Schmein–Gordon equation admits both real- and complex-valued planewave solutions. And in both cases, the dispersion relation of its planewave solutions differs only slightly from that of the Klein–Gordon equation (Eqn. 7, p. 124).

$$\omega = \sqrt{-\kappa^2 + m_0^2} \tag{13}$$

Now, although this dispersion relation defines ω to be real-valued only when $\kappa^2 \leq m_0^2$, the Schmein–Gordon equation still admits complex-valued solutions ψ of the form of Equation 5 when $\kappa^2 > m_0^2$. But in that case, the solutions are not planewaves, but rather exponentials. This is illustrated in Figure 15-ii. Nonetheless, the Schmein–Gordon equation does admit planewave solutions when $\kappa^2 \leq m_0^2$. And in that case, the Schmein–Gordon equation also admits real-valued planewave solutions $\psi_{\Re}$ of the form of Equation 8. This is illustrated in Figure 15-iii.

Once again, $\langle P, \mathcal{T} \rangle$ admits permissible coodinatizations according to which $f_{\phi'}$ solves the Schmein–Gordon equation. However, the coordinate transformations involved are not quite as simple as they were in earlier examples. Consider for example the move from Figure 15-i to 15-iii. The former illustrates a real-valued solution to the Klein–Gordon for which $\kappa^2 > m_0^2$, and given the preceding comments, this solution would therefore require

some adjustment to the mass or wavenumber before it could solve the Schmein–Gordon equation. Thus whereas it was shown above that the autohomeormorphism relating $\psi = f_\phi$ to some $f_{\phi'}$ that solves the Squaringer or free wave equation was a simple stretching of the t-axis, the autohomeomorphism relating $\psi = f_\phi$ to some $f_{\phi'}$ that solves the Schmein–Gordon equation will sometimes involve an additional dilation of spatial coordinates along the direction of the wavevector, which would affect the wavenumber accordingly.[17] But of course such coordinate transformations are still autohomeomorphisms of the standard topology of $\mathbb{R}^4$, and so $\langle P, \mathcal{T} \rangle$ will admit coordinatizations according to which $f_{\phi'}$ solves the Schmein–Gordon equation.

Given this, it is worth enquiring after the symmetry group of the Schmein–Gordon equation. Unlike the Klein–Gordon equation, the Schmein–Gordon equation is not Lorentz-covariant, meaning that the Schmein–Gordon equation is a rival systematization of $\langle P, \mathcal{T} \rangle$ upon the symmetries of which would supervene some non-Minkowskian geometrical structure.

5.3 Considering More General Solutions

It has been shown above that, as expected, both complex- and real-valued planewave solutions to the Klein–Gordon equation, when considered 'up to homeomorphism' as $\langle P, \mathcal{T} \rangle$, admit permissible coordinatizations ϕ' according to which the $f_{\phi'}$ solve other, equally simple and strong equations with different covariance groups. If some such $\langle P, \mathcal{T} \rangle$ were being taken seriously as a potential example of the primitive ontology of a constructivist project, it would be necessary to outline the full range of rival systematizations, find a way to compare their simplicity and strength, and then defend a Lorentz-covariant one as the 'liberalized', best-systems dynamical law. But as noted above, a single planewave solution to the Klein–Gordon equation constitutes an extremely simple and symmetric space of field points, and one ought not be surprised if such a

[17] Or, instead of resolving the relative magnitude of κ^2 and m_0^2 by a change in spatial coordinates, the same could be achieved by an appropriate change in the value of m_0. But having the freedom to make changes to m_0 would make it a supervenient quantity. This is discussed further below.

$\langle P, \mathcal{T} \rangle$ cannot support a best-systems account of Minkowski geometrical structure.[18]

Instead, the following question should be addressed: given that planewaves of the form of Equations 5 and 8 are not the most general solutions to Klein–Gordon equation, might its solutions of a more general form have characteristics such that the corresponding $\langle P, \mathcal{T} \rangle$ admit no permissible recoordinatizations such that the $f_{\phi'}$ solve even the few rival equations introduced above? That is, might some of the competition be eliminated by choosing a $\langle P, \mathcal{T} \rangle$ that comprises the point-like parts of a more complicated solution? This section addresses that question in two ways: first, by considering more general solutions to the Klein–Gordon equation as introduced above; second, by considering the solutions to a more general form of the Klein–Gordon equation itself.

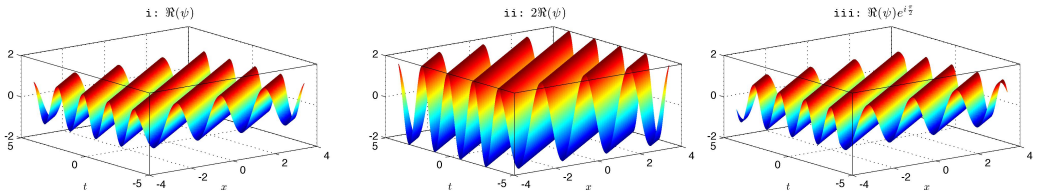
In both cases, it will be shown that by considering just slightly more complicated solutions, it becomes much more difficult to find permissible coordinatizations with respect to which the $f_{\phi'}$ solve any other equations at all. Thus the onus is on the critic of the dynamical approach to show that some other non-Lorentz-covariant equation deserves the title of 'dynamical law' for the systems under consideration. These examples are still extremely simple in comparison with anything that could potentially describe a more realistic, specially relativistic world, but this just goes to bolster the plausibility of the constructivist's best-systems claim.

Linear Combinations of Planewave Solutions

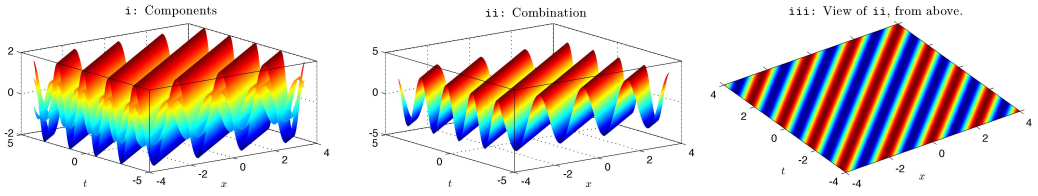
It was reported in §5.1 that the Schrödinger and Klein–Gordon equations both admit complex-valued planewave solutions ψ of the form of Equation 5. Being linear equations, they also admit solutions of the form of any linear combination of its planewave solutions. That is, any solution of the form

$$\psi = \sum_{\kappa} A_{\kappa} e^{i(\bar{\kappa} \cdot \bar{x} - \omega_{\kappa} t)}$$

[18] Indeed, for a single, real-valued planewave solution $\psi_{\Re}$, certain permissible coordinatizations of $\langle P, \mathcal{T} \rangle$ will yield no change in $f_{\phi'}$ along particular coordinates, allowing for systematizations of fewer variables. For example, under a recoordinatization ϕ' according to which $\partial_t f_{\phi'} = 0$, it turns out that $f_{\phi'}$ solves the Helmholtz equation: $\nabla^2 \psi = -m_0^2 \psi$. However, this will no longer be the case for more general solutions to the Klein–Gordon equation. (*cf.* fn. 19, p. 138.)



(a) Three planewaves of equal $\bar{\kappa}$, ω, and thus $\bar{v}_p$.



(b) Sum of the three planewaves above.

Figure 16: The linear combination of three real-valued, planewave solutions to the Klein–Gordon equation ($\omega = \sqrt{\kappa^2 + m_0^2}$), all with equal wavevectors $\bar{\kappa}$, angular frequencies ω, and thus phase velocities $\bar{v}_p$. The summands are shown in 16(a): plot i has an amplitude of $A = 1$, and therefore matches Figure 12-i; plot ii differs from i by having a doubled amplitude, $A = 2$; and plot iii differs from plot i by a phase shift of $\frac{\pi}{2}$. Their sum is shown in 16(b): plot i shows the summands superimposed; plot ii shows their sum; plot iii shows their sum from above. As in earlier figures, $m_0 = \hbar = 1$ and $\kappa = 3$.

will solve the Schrödinger equation (Eqn. 3), so long as the dispersion relation $\omega_\kappa = \kappa^2/2m$ holds for each of the summands. And of course, the same could be said for the Klein–Gordon equation, but with two changes: first, the dispersion relation that would have to hold for each of the summands is $\omega_\kappa = \sqrt{\kappa^2 + m_0^2}$; second, the linear combination claim would also hold for real-valued planewave solutions $\psi_\Re$.

What are the salient characteristics of linear combinations of real-valued planewaves? As shown below, linear combinations of planewaves sometimes constitute planewaves themselves, and at other times result in more complicated wave configurations. Consider first an example of the former: a linear combination of different planewaves, all of which have the same wavenumber κ, angular frequency ω, and therefore phase velocity $\bar{v}_p$, the sum of which is another planewave. Figure 16(a) illustrates three such planewave summands. From the discussion above, it should be clear by inspection that each of the summands (i–iii) is homeomorphic to a solution to any of the three rival equations introduced above. The three planewaves' sum is illustrated in Figure 16(b). There too, it

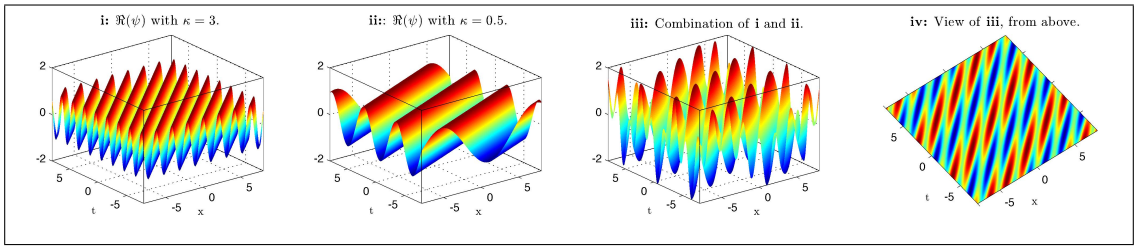


Figure 17: The sum of two real-valued planewave solutions to the Klein–Gordon equation, for which $\omega = \sqrt{\kappa^2 + m_0^2}$. As above, $A = m = \hbar = 1$, so that plot i matches Figure 12-i again. The difference in the summands' wavenumbers κ ensures that their sum is not a planewave.

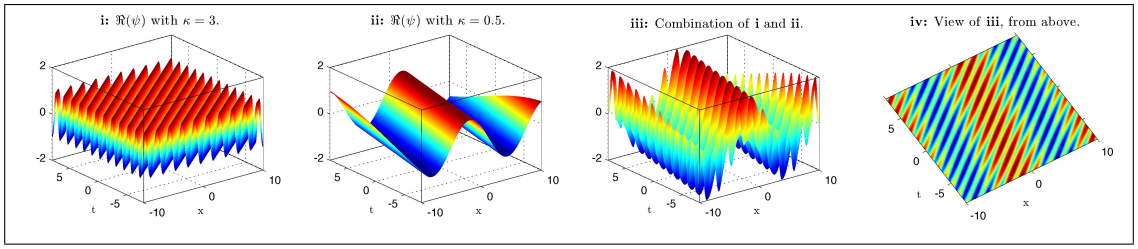


Figure 18: The sum of two real-valued planewave solutions to the Squaringer equation, for which $\omega = \kappa^2/2m$. As above, $A = m_0 = \hbar = 1$. Note that plots i and ii have the same value for κ as the corresponding plots in Figure 17.

should be clear by inspection that the resulting planewave is also homeomorphic to a solution to any of the same rival equations. From this it follows that a linear combination of planewave solutions to the Klein–Gordon equation, all having equal wave vectors and angular frequencies, results in another solution to the Klein–Gordon equation that, when taken up to homeomorphism as $\langle P, \mathcal{T} \rangle$, admits alternative coordinate-dependent descriptions $f_{\phi'}$ from among which can be found solutions to any of the rival equations introduced above.

But now consider a linear combination of planewave solutions to the Klein–Gordon equation, the wavenumbers (and thus phase velocities) of which differ from one another, and the sum of which is therefore not a planewave. Figure 17 illustrates two such planewaves.[19] As before, it is clear by inspection that both of the summands (i, ii) are homeomorphic to a solution to any of the rival equations introduced above. However, it is less clear whether their sum (iii, iv), which is another real-valued solution to the Klein–Gordon equation, is also homeomorphic to solutions of those rival equations. In

[19] Note that for such solutions, the systematization mentioned in footnote 18 (p. 136) is no longer possible, as there are no dimensions along which the field values are constant.

fact it is. To see this for the Squaringer equation, consider what would result from the following process: first, consider two planewave solutions to the Squaringer equation with wave-vectors equal to those of the planewave summands illustrated in Figure 17; then, consider the linear combination of those planewaves. This linear combination is illustrated in Figure 18. By the fact that the summands (i, ii) solve the Squaringer equation, it follows that the sum (iii, iv) does as well. Then, comparing the sum to Figure 17-iv reveals that the two are indeed homeomorphic—i.e. their differences can be accommodated by some permissible autohomeomorphism φ. This is illustrated and explained in Figure 19. A similar exercise can be carried out for linear combinations of two planewave solutions to the free wave or Schmein–Gordon equations.

From this it follows that a linear combination of two real-valued planewave solutions to the Klein–Gordon equation is no more viable as a supportive example of the constructivist's account of Minkowski geometry than a single planewave solution was. However, things change quite quickly when the number of summands increases. Consider the situation illustrated in Figure 20, which shows the linear combination of three planewave solutions to the Klein–Gordon equation, all having distinct wavenumbers. Is the sum homeomorphic to a solution to any of the three rival equations introduced above? If it is, those solutions can not be found by the process above, for the sum in Figure 20-iv is not obviously homeomorphic to what results from the preceding operation of taking planewave solutions to, say, the Squaringer equation, all having corresponding values for κ, and then taking their linear combination. The result of that process is illustrated in Figure 21. To see what the trouble is, consider the steps that would be involved in repeating the process that was outlined in Figure 19. After using a time dilation to transform the first planewave summand into a solution to the Squaringer equation, it would once again be necessary to boost into the 'rest frame' of that planewave before using another time dilation to transform the second summand into a solution to the Squaringer equation. But when this process is attempted for the third summand, trouble arises. There is no mutual 'rest frame' for the first and second summands, from which

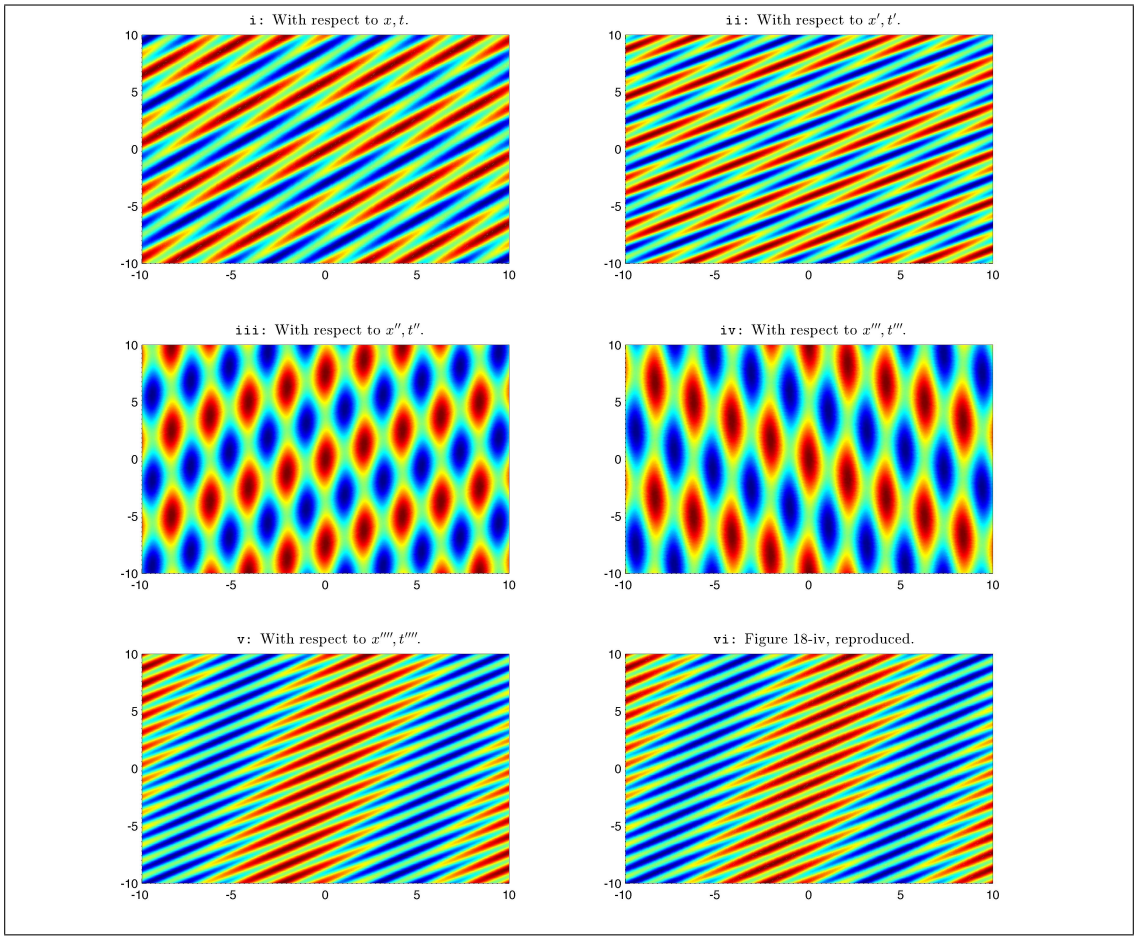


Figure 19: The relationship between plots 17-iv and 18-iv. Plot i shows the former, and plots ii–v show a series of transformations under which plot i comes to solve the Squaringer equation. That is, plots ii–iv show a series of transformations under which Figure 17-iv becomes Figure 18-iv. In reviewing these steps, it will be helpful to think of plot i in terms of the two planewave summands illustrated in Figure 17. To begin, plot ii shows the result of the time dilation, mentioned earlier (p. 128), under which the first planewave summand shown in Figure 17-i would come to satisfy the Squaringer equation. For this step, $x \to x' = x$ and $t \to t' = t\sqrt{\kappa^2 + m_0^2}/(\kappa^2/2m)$, with $\kappa = 3$. Now that one of the summands of Figure 17 has been transformed into a solution to the Squaringer equation, the following three steps must do the same for the second summand of Figure 17, but without ruining the work accomplished in this first step. To make this possible, plot iii shows the result of boosting into the 'rest frame' of the now-dilated first planewave summand. For this step, $t' \to t'' = t'$, and $x' \to x'' = x' - v_d t'$, where $v_d = 1.5$ is the phase velocity of the first planewave summand after the aforementioned time dilation. Being now in the 'rest frame' of the first planewave summand, another time dilation can safely be carried out to transform the second planewave summand into a solution to the Squaringer equation. For this step, $x'' \to x''' = x''$, and $t'' \to t''' = -1.3454t''$. (The decimal value can be solved for by requiring that the final result of these steps be a solution to the Squaringer equation.) Finally, plot v shows the result of boosting back by an equal and opposite to the previous boost, so as to undo the changes brought about in moving from plot ii to plot iii. For this step, $t''' \to t'''' = t'''$, and $x''' \to x'''' = x''' + v_d t'''$. Plugging in values for κ, ω, and v_d, these coordinate transformations combine to the overall transformation according to which $x \to x'''' = x + 2.6149t''''$, and $t \to t'''' = -0.9454t$. Thus the exponent of the first planewave summand of Figure 17 transforms as $\kappa x - \omega t = 3x - \sqrt{10}t \;\longrightarrow\; 3(x'''' - 2.6149t'''') - \sqrt{10}(-1.0577)t''''$. This simplifies to $\kappa x - \omega t$, with $\kappa = 3$ and $\omega = 4.5 = \kappa^2/2m$, thereby satisfying the dispersion relation of the Squaringer equation (p. 131). Similarly, the exponent of the second planewave summand of Figure 17 transforms as $0.5x - \sqrt{1.25}t \;\longrightarrow\; 0.5(x'''' - 2.6149t'''') - \sqrt{1.25}(-1.0577)t''''$. This simplifies to $\kappa x - \omega t$, with $\kappa = 0.5$ and $\omega = 0.125 = \kappa^2/2m$, also satisfying the dispersion relation of the Squaringer equation. The net effect of these transformations on the planewaves' sum can be seen by comparing plots v and vi.

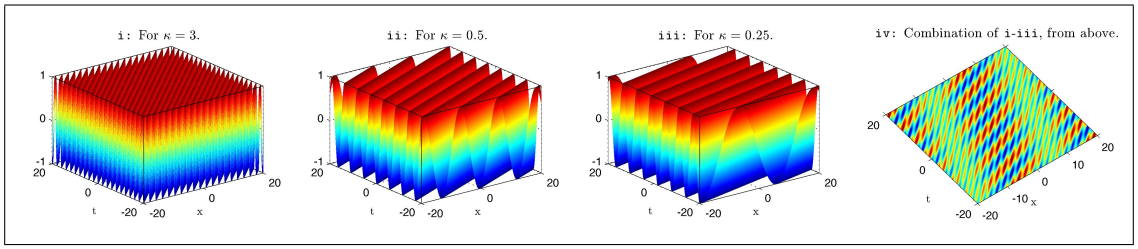


Figure 20: The linear combination of three planewave solutions to the Klein–Gordon equation, for which $\omega = \sqrt{\kappa^2 + m_0^2}$. As above, $A = m = \hbar = 1$. Plots i and ii match those of Figure 17 in a different scale.

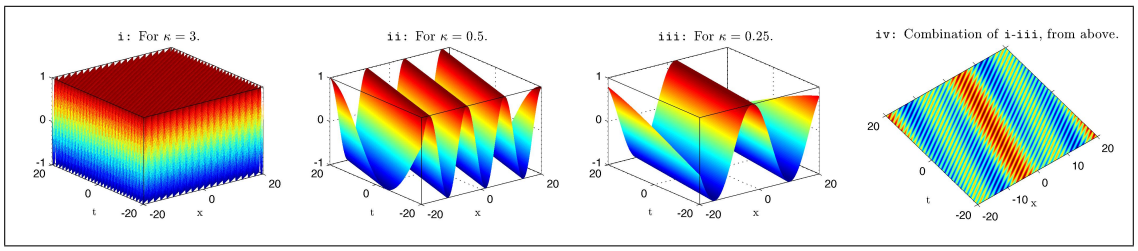


Figure 21: The linear combination of three planewave solutions to the Squaringer equation, for which $\omega = \kappa^2/2m$. As above, $A = m = \hbar = 1$. Plots i and ii match those of Figure 18 in a different scale.

one could carry out a time dilation on the third summand without affecting what was accomplished in the previous two operations. The problem is simply that there are now more planewaves than there are dimensions. Any time dilation that transforms the third planewave summand into a solution of the Squaringer equation will interfere with the dispersion relation of at least one of the first two planewave summands.[20]

Of course, it should be granted that this only shows that by taking up to homeomorphism a linear combination of a sufficient number of real-valued planewave solutions to the Klein–Gordon equation, all differing in wavenumber, there is no combination of boosts (shearing transformations) and dilations (stretching transformations) under which $\langle P, \mathcal{T} \rangle$ would be recoordinatized so that $f_{\phi'}$ solves one of the rival systematizations outlined above. Whether some other autohomeomorphism of $\mathcal{R}^4$ might do the trick, or whether there may be alternative coordinatizations ϕ' according to which the $f_{\phi'}$ solve some other equations altogether, is an open question. But for the time being, the lesson is simply that by taking a linear combination of planewave solutions to the Klein–Gordon equation

[20] Note that this claim is made in regard to two-dimensional examples. In taking a four-dimensional topological structure as primitive, this would only be true in the case of *five* or more planewave solutions to the Klein–Gordon equation, all differing in wavenumber. A similar comment applies to the claim made in footnote 19 (p. 138).

with different wavenumbers, the obvious competition for the best-systems dynamical law of $\langle P, \mathcal{T} \rangle$ at least diminishes. Thus the onus is on the critic of the dynamical approach to find an autohomeomorphism of $\mathcal{R}^4$ such that the corresponding $f_{\phi'}$ would solve some otherwise-covariant equation, the simplicity and strength of which could rival those of the Klein–Gordon equation.

Solutions to the General Klein–Gordon Equation

So far in this chapter, the conversation has been restricted to solutions of the Klein–Gordon equation as introduced in §5.1. But we might also consider solutions to the more general form of the Klein–Gordon equation. The general Klein–Gordon equation is introduced below, and a similar case-study is carried out with respect to its solutions. Although the equations under consideration will be a bit more complicated, the overall message is the same—namely, that while the general Klein–Gordon equation's simplest solutions can indeed be recoordinatized so as to satisfy some of the (generalized) rival systematizations introduced above, only slightly more complicated examples are necessary to make it much more difficult to find any coordinatizations ϕ' with respect to which the corresponding $f_{\phi'}$ would solve any other equations.

In particular, this is shown in two ways. First, it is shown that as the solutions to the general Klein–Gordon become more complicated than simple planewaves, finding a rival systematization will sometimes require that more and more of the system under consideration be taken to supervene upon the 'liberalized' best-system dynamical laws. For instance, it is shown that taking mass as a supervenient quantity is the only way for certain solutions of the general Klein–Gordon equation to be systematized by the more general form of the Schmein–Gordon equation. It is also shown that in certain cases, the potential fields themselves would need to be taken as supervenient. The second way in which it becomes more difficult to find rival systematizations for solutions to the general Klein–Gordon equation comes up in the range of its solutions' values. Whereas the solutions considered above all had values in the range of $-A \leq \psi \leq A$, some solutions

to the general Klein–Gordon equation take a more complicated, exponential form. This brings about the added difficulty of finding solutions to rival systematizations that even comprise the same set of field-values, let alone the same topological structure on those values.

To begin, the general Klein–Gordon equation must be introduced. Above, the 'free' Klein–Gordon equation was given a heuristic derivation from the relativistic energy–momentum relation for a free particle (§5.1). The more general form of the Klein–Gordon equation can be given a similar derivation from the relativistic energy–momentum relation for a particle in the region of some external potential.[21] Let the potential be $\Phi = (S, \bar{V})$, so that $E \to E - S$ and $\bar{p} \to \bar{p} - \bar{V}$. Thus the relativistic energy–momentum relation is changed.

$$E^2 = p^2 + m_0^2 \quad \to \quad (E - S)^2 = (\bar{p} - \bar{V})^2 + m_0^2$$

As before, energy and momentum can be turned into operators via the standard correspondence relations $E \to i\partial_t$ and $p \to -i\nabla$. This gives the general Klein–Gordon equation.

$$(i\partial_t - S)^2\psi - (-i\nabla - \bar{V})^2\psi = m_0^2\psi$$

For the sake of comparison with the 'free' Klein–Gordon equation (Eqn. 4, p. 123), the general Klein–Gordon equation can be put into the following form,

$$-\partial_t^2\psi + \nabla^2\psi + U\psi = m_0^2\psi \tag{14}$$

where $U(t, \bar{x})$ has both vector and scalar components.

$$U = \underbrace{-i\partial_t S - i\nabla\bar{V} - iS\partial_t - i\bar{V}\nabla}_{\text{vector}} + \underbrace{S^2 - \bar{V}^2}_{\text{scalar}}$$

Before going any further, a few things should be said about how to understand potentials in the context of the constructivist project. In one scenario, the potentials under consideration could be understood as additional physical fields, interacting with

[21] The following discussion draws in part from unpublished resources made available by Elster [2013].

the scalar-valued solutions to the Klein–Gordon equation, as part of the package to be systematized. This would entail making an addition to the constructivist's fundamental ontology, and some remarks about making such changes will be given in Chapter 6. In that case, one would not have the freedom to make changes to the potential terms when conjuring up rival systematizations of the field values; the potentials could only be recoordinatized in the same ways as the field-points of the solution to the general Klein–Gordon equation. But in another scenario, one might allow the constructivist the freedom of making changes to the potential terms in looking for rival systematizations of the field values. This would entail a somewhat bolder supervenience claim, for it would suggest that the potentials themselves supervene upon the 'liberalized' best-systems dynamical laws. The differences between these two approaches should become clear in the following examples.

With regard to the solutions of the general Klein–Gordon equation (Eqn. 14), it should be noted from the outset that while it does admit complex-valued solutions ψ in the form of Equation 5, the only circumstances under which it admits real-valued solutions $\psi_{\Re}$ of the form of Equation 8 is when $\Phi = (S, \bar{V}) = 0$. To see this, note that the final two of the four vector components of $U(t, \bar{x})$ are complex. For the same reasons rehearsed in the introduction to the Squaringer equation (p. 131), their presence would prevent real-valued solutions.[22] And so, if this case-study were limited to real-valued examples, the general Klein–Gordon equation would be of no use in the search for a supportive example of the constructivist project. Indeed, the only generalization of the Klein–Gordon equation to admit real-valued solutions is the addition of a so-called 'Lorentz scalar'. While not a true interaction term, this is a scalar potential that adds to the mass term, giving effectively another version of the Klein–Gordon equation.

$$-\partial_t^2\psi + \nabla^2\psi = (m_0^2 + L)\psi$$

[22]The exception, of course, would be if the vector components of $U(t, \bar{x})$ were to cancel out, so that $(-i\partial_t S - i\nabla\bar{V} - iS\partial_t - i\bar{V}\nabla)\psi_{\Re} = 0$. But the end result would be a mere Lorentz scalar potential, described at the end of this paragraph.

This equation admits both complex- and real-valued solutions of the form of Equations 5 and 8. For both, the dispersion relation differs only slightly from that of the free Klein–Gordon equation.

$$\omega = \sqrt{\kappa^2 + m_0^2 + L}$$

But as has been shown above, this kind of change in the dispersion relation amounts to no difference at the level of topology. Thus solutions to the Klein–Gordon equation with a Lorentz scalar potential are of no advantage in the present project. Instead, the following discussion will be held in the context of complex-valued solutions of the general Klein–Gordon equation, meaning that the rest of this section will be concerned with topological spaces of complex values, $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$. Still, because each of the rival systematizations introduced above also admit complex-valued solutions, the rival systematizations under consideration will be the generalized versions of the same equations.

The form of the complex-valued solutions to the general Klein–Gordon equation hinges upon the choice of potential Φ. It will be helpful to begin with a fairly simple example. If a potential is chosen for which $\partial_t S = \nabla \bar{V} = 0$, then the general Klein–Gordon equation can again be expressed in terms of $U(t, \bar{x})$, this time with the latter taking a simpler form.[23]

$$U = -2iS\partial_t - 2i\bar{V}\nabla + S^2 - \bar{V}^2$$

And as usual, the differences between solutions to the general Klein–Gordon equation and those of the free Klein–Gordon equation will be encoded in the equations' respective dispersion relations. As in Equation 7 (p. 124), that of the free Klein–Gordon equation was $\omega = \sqrt{\kappa^2 + m_0^2}$, whereas that of the general Klein–Gordon equation now takes the following, slightly more complicated form.

$$\omega = \sqrt{(\bar{\kappa} - \bar{V})^2 + m_0^2} + S \tag{15}$$

Now, if the potential Φ were chosen to satisfy the condition $\partial_t S = \nabla \bar{V} = 0$ by S and

[23] Note that the factor of 2 in the first two terms comes from applying the product rule to the first and second terms of $U(t, \bar{x})$ in Equation 14.

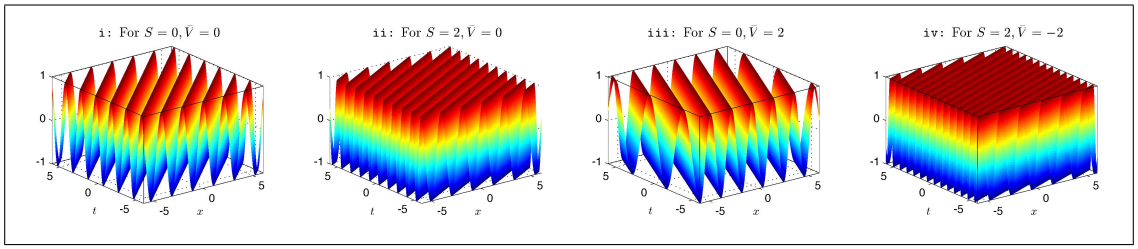


Figure 22: Real-valued components of planewave solutions to the general Klein–Gordon equation, when Φ is constant. As above, $A = m = \hbar = 1$, and $\kappa = 3$. Plot i shows the solution for no potential, and therefore matches Figure 12-i (p. 128). Plots ii–iv show planewave solutions for other constant potentials.

$\bar{V}$ being entirely constant, then S and $\bar{V}$ would affect the solution's dispersion relation in a relatively straightforward way: for a given wavenumber κ, the angular frequency ω will increase proportionally with S, and nearly so with $\bar{V}$.[24] And as illustrated in Figure 22, solutions to the general Klein–Gordon equation with constant potential are still planewave solutions (and their linear combinations). Therefore, like solutions to the general Klein–Gordon equation with a Lorentz scalar potential introduced above, these solutions to the general Klein–Gordon equation with a constant potential are of no advantage to the constructivist's project.

But of course the requirement that $\partial_t S = \nabla \bar{V} = 0$, which was chosen above merely for simplicity's sake, does not rule out all non-constant potentials. In fact, as illustrated in Figure 23, solutions to the general Klein–Gordon equation with non-constant potentials still satisfying this constraint are in fact not planewaves. Many, however, such as those illustrated in plots ii–iv, are clearly homeomorphic to planewaves, and would therefore admit recoordinatizations ϕ' with respect to which the $f_{\phi'}$ solve the rival equations above. But some, such as the solution illustrated in Figure 23-i, have lost homeomorphicity with a planewave solution. This is due to the asymptotes at values of x for which S causes $\omega \to 0$. If this solution is taken up to homeomorphism as $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$, does it admit any permissible recoordinatizations with respect to which the $f_{\phi'}$ solve any of the rival systematizations introduced above? For instance, we might ask this with respect to the

[24] For certain values of $\bar{V}$, which depend upon that of κ, the relationship is inversely proportional. But in any case, the effect on ω of a change in $\bar{V}$ is diminished in proportion to the value of m_0.

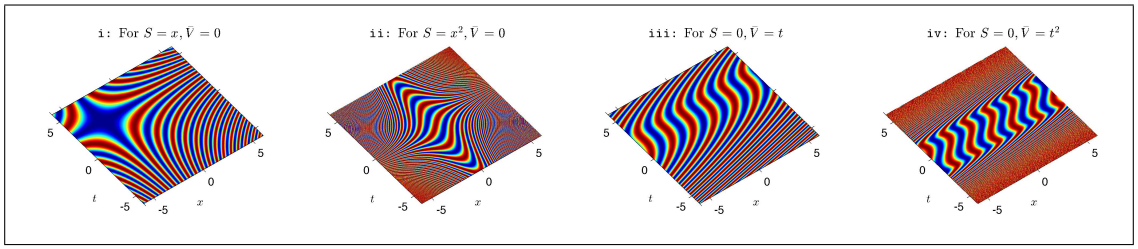


Figure 23: Real-valued components of solutions to the general Klein–Gordon equation with simple but non-constant potentials, for which $\partial_t S = \nabla \bar{V} = 0$. As above, $A = m_0 = \hbar = 1$, and $\kappa = 3$. (See also the caption to Figure 26.)

generalized form of the Schmein–Gordon equation,

$$\partial_t^2 \psi + \nabla^2 \psi + U\psi = -m_0^2 \psi \tag{16}$$

which can be derived by the same procedure as for the generalized Klein–Gordon equation (p. 143). The same restriction that $\partial_t S = \nabla \bar{V} = 0$ gives $U(t, \bar{x})$ a similar form,

$$U = 2iS\partial_t - 2i\bar{V}\nabla - S^2 - \bar{V}^2$$

and Equation 16 admits complex-valued solutions of the form of Equation 5, having a slightly different dispersion relation.

$$\omega = \sqrt{-(\bar{\kappa} - \bar{V})^2 + m_0^2} + S$$

And, as with the 'free' Schmein–Gordon equation (Eqn. 12, p. 134), there is a requirement on the relationship between m_0 and κ in order to avoid a complex-valued dispersion relation, which would give rise to dampening terms in the solution. This time, the requirement is that $(\bar{\kappa} - \bar{V})^2 \leq m_0^2$.

The question at hand is this: could the solution to the general Klein–Gordon equation illustrated in Figure 23-i, when taken up to homeomorphism as $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$, be recoordinatized so as to solve the general Schmein–Gordon equation? From a glance at Figure 24-i, it would appear that it cannot. At least, for the same choice of potential terms $\bar{V}$ and S, and for the same values of A, κ, and m_0, the aforementioned requirement that $(\bar{\kappa} - \bar{V})^2 \leq m_0^2$ is not satisfied. As a result, the values of the solution to the general Schmein–Gordon equation are dampened. This results in an entirely different set of values than those

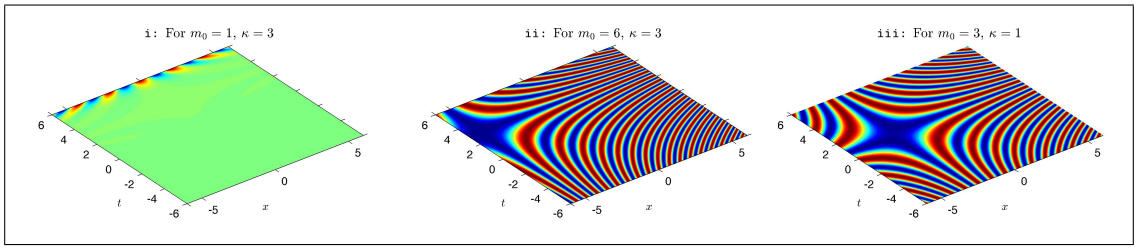


Figure 24: Real-valued components of solutions to the general Schmein–Gordon equation, with different values of m_0 and κ. Otherwise, the values are all as above: $A = \hbar = 1$. For each plot, the potential Φ has the same components as in Figure 23-i: $S = x$ and $\bar{V} = 0$.

illustrated in Figure 23-i, and no amount of recoordinatizing them will result in the solution illustrated there. But of course, in looking for rival systematizations of $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$, there are no rules against considering solutions to other equations with a different value for κ. Taking this idea a bit further, one might even allow oneself the liberty of considering alternative values for m_0. Doing so would make mass supervene upon the 'liberalized' best-system dynamical laws, and in this case it would facilitate meeting the requirement on κ and m_0 for the Schmein–Gordon equation.[25] Some solutions to the generalized Schmein–Gordon equation, with the same choice of potential as Figure 23-i, but different values for m_0 and κ, are illustrated in Figure 24-ii,iii. It is clear by inspection that for the right choices of m_0 and κ, the solutions to the general Schmein–Gordon equation are indeed homeomorphic to those of the Klein–Gordon equation, even of the type illustrated in Figure 23-i. However, it should not go unnoticed that even in the case of a very simple potential, finding a way to systematize $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$ with the Schmein–Gordon equation has required that mass be understood as a supervenient quantity.

Even more must be taken as supervenient when other potentials are considered. It was mentioned above that in exploring rival systematizations of some given $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$, one might or might not consider systematizations involving different potentials. In finding a rival systematization for the solution illustrated in Figure 23-i, no changes were made to the potential, but only to κ and m_0. However, it should be noted that the reason these

[25] Note that something similar was going on when the free wave equation was listed as a rival systematization of a planewave solution to the Klein–Gordon equation, since the free wave equation is effectively the Klein–Gordon equation with $m_0 = 0$. See also the earlier discussion on changing the value of m_0, in footnote 17 on page 135.

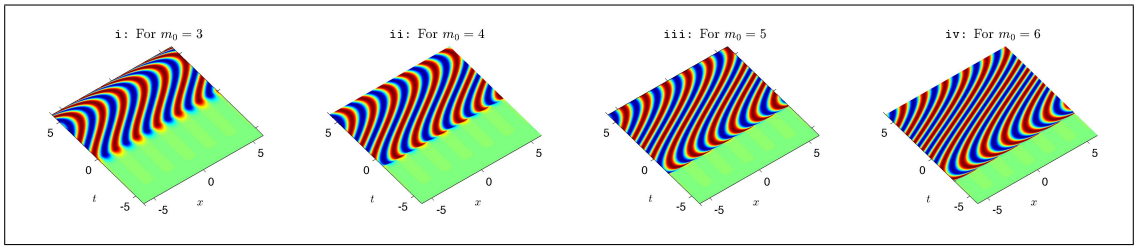


Figure 25: Real-valued components of solutions to the general Schmein–Gordon equation, with different values of m_0. Otherwise, the values are all as in most illustrations above: $A = \hbar = 1$, and $\kappa = 3$. For the solutions in each plot, the potential Φ has the same components as in Figure 23-iii: $S = 0$ and $\bar{V} = t$.

relatively minor changes were sufficient in finding a homeomorphic solution to the general Schmein–Gordon equation was that the vector potential $\bar{V}$ had been chosen to be zero. For whenever $\bar{V}$ is constant for a solution to the general form of the Schmein–Gordon equation, the aforementioned requirement that $(\bar{\kappa} - \bar{V})^2 \leq m_0^2$ can be satisfied by an appropriate choice of m_0 or κ. However, were $\bar{V}$ to vary with respect to t—and note that this would still satisfy the earlier simplifying condition of $\partial_t S = \nabla \bar{V} = 0$—then at certain values of t, no choice of (constant) m_0 or κ would suffice. In other words, the same trick of hand-picking the value for m_0 and κ would not facilitate finding a solution to the general Schmein–Gordon equation that is homeomorphic to the solution illustrated in Figure 23-iii, where $S = 0$ and $\bar{V} = t$. And indeed, as shown in Figure 25, if the same values are chosen for A, κ, $\bar{V}$, and S, then the appearance of complex-valued dispersion relations in the solution to the general Schmein–Gordon equation causes the field values to go to zero at certain values of t.

Now, it should be re-stated that the solution to the general Klein–Gordon equation illustrated in Figure 23-iii is of no use to the constructivist's project. It is homeomorphic to a planewave, and is therefore homeomorphic to solutions to the rival systematizations outlined above. That is to say, it allows permissible recoordinatizations ϕ' with respect to which the $f_{\phi'}$ solve those rival systematizations. However, it serves as a useful example in showing that for any solution to the general Klein–Gordon equation with a non-constant vector potential, there are serious obstacles in finding any recoordinatization with respect to which it could solve an equation like the general Schmein–Gordon equation. In particular, it shows that the only solutions to the general Schmein–Gordon equation

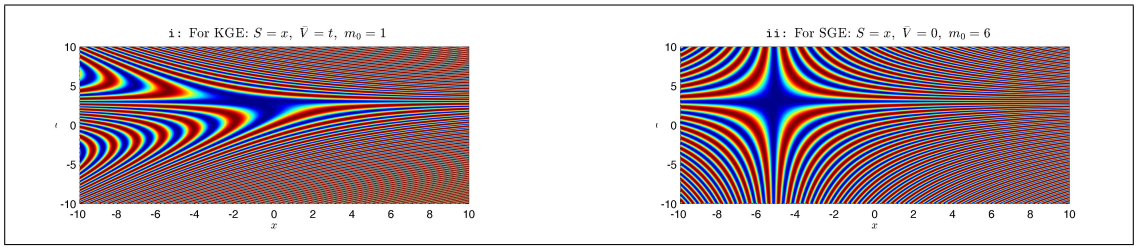


Figure 26: Real-valued components of solutions to (i) the general Klein–Gordon equation, and (ii) the general Schmein–Gordon equation. As above, $A = \hbar = 1$, and $\kappa = 3$. Plot i shows a solution which differs from that illustrated in Figure 23-i only with regard to the vector component $\bar{V}$ of the potential. Plot ii shows the same solution illustrated in Figure 24-ii, but on a different scale. Because the vector component $\bar{V}$ of the potential chosen for the solution illustrated in Plot i brings about no changes to the topological structure of the field values, that solution it is still homeomorphic to the solution illustrated in Plot ii.

that will be homeomorphic to it will not have a vector potential $\bar{V}$ that varies along t in the same way. Rather, this particular $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$ can only be recoordinatized so as to solve the general Schmein–Gordon equation with a different potential. Indeed, if allowed the freedom of hand-picking the scalar and vector components of the potential, it is not difficult to find an example. For given that the solution to the general Klein–Gordon equation illustrated in Figure 23-iii is homeomorphic to a planewave solution, the choice of $S = \bar{V} = 0$ for a (planewave) solution to the general Schmein–Gordon equation would suffice! Or, even if a non-constant vector component like $\bar{V} = t$ (as in Figure 23-iii) were chosen in conjunction with a non-constant scalar component like $S = x$ (as in Figure 23-i) for the Klein–Gordon equation, the result is still homeomorphic to the solution illustrated in Figures 24-ii,iii. This is explained further in the caption to Figure 26, but the upshot is this: for some solutions to the general Klein–Gordon equation, having relatively simple potential terms, rival systematizations can be found in other particular equations only when the potentials themselves are taken to supervene upon the 'liberalized' best-systems dynamical laws. Granted, other rival systematizations might be found which do not require that the potentials be taken to supervene. Finding them, however, will be more difficult than exploring the solutions to now-familiar equations like the Schmein–Gordon equation.

The preceding examples show that by considering solutions to the general Klein–Gordon equation with potential terms, finding rival systematizations quickly begins to

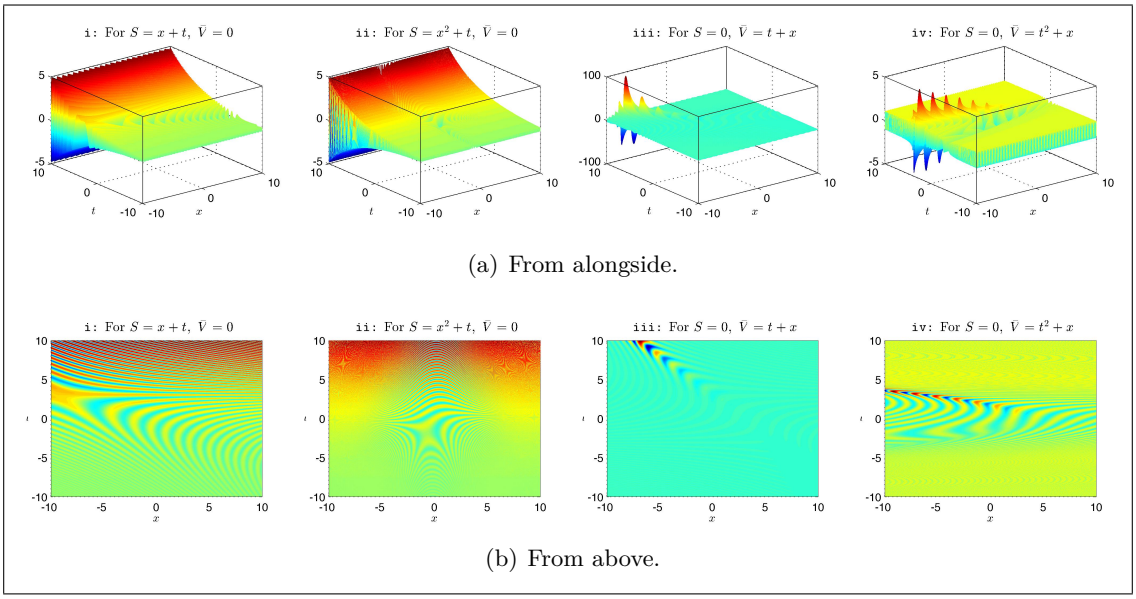


Figure 27: Real-valued components of complex-valued solutions for the general Klein–Gordon equation, when $\partial_t S \neq 0$ or $\nabla \bar{V} \neq 0$. As above, $A = m_0 = \hbar = 1$, and $\kappa = 3$. However, regarding the value of A, note the difference in the range of field values. The solutions are illustrated (a) from alongside and (b) from above. Plots i–iv should be compared with the respective plots in Figure 23 (p. 147), as the difference in potentials illustrates the dampening and amplifying effects of allowing $\partial_t S \neq 0$ or $\nabla \bar{V} \neq 0$.

require that more of the system under consideration be construed as supervenient. Yet another difficulty in finding rival systematizations of solutions to the general Klein–Gordon equation can be illustrated if the potentials are made just slightly more complicated. For all of the above has been said with respect to solutions to a relatively simple form of the general Klein–Gordon equation—namely, that for which $\partial_t S = \nabla \bar{V} = 0$. For more general potentials, $U(t, \bar{x})$ takes on a more complicated form, and so the solutions to Equation 14 (p. 143) inherit a correspondingly more complicated dispersion relation.

$$\omega = \sqrt{(\bar{\kappa} - \bar{V})^2 + m_0^2 + i\partial_t S + i\nabla \bar{V}} + S$$

The most significant change in the field values of these solutions comes from the fact that the dispersion relation's imaginary-valued terms act as amplifying and dampening terms on the real-valued component of the solution, resulting in the kind of behavior illustrated in Figure 27. The potentials under consideration are still extremely simple, but already they reveal features that would decrease the likelihood of finding an alternative systematization. For now a very different set of field-points are under consideration; the

values do not all land somewhere in the range of $-A \leq \psi \leq A$. And even if it could be shown that some solution to another, rival systematization comprised the same set of field values, it would be especially difficult to find one for which those values were given the same topological structure. The reason for this is simply that in addition to having a dampening and amplifying effect, the imaginary-valued terms in the equation's dispersion relation also bring about a distorting effect on the field-points. This can be seen by comparing the additional differences between the corresponding plots in Figures 27(b) and 23.

5.4 Summary and Reflections

The goal of this chapter has been to use solutions to the Klein–Gordon equation as a case study in the constructivist's best-systems account of Minkowski spacetime structure. That is to say, solutions to the Klein–Gordon equation have been the focus in an exploration of the salient characteristics of a topologically structured set of scalar values that would be systematized better (i.e. with some better balance of simplicity and strength) by a Lorentz-covariant equation than by any equation with another covariance group.

First, it was shown that the simplest, planewave solutions to the Klein–Gordon equation can indeed be recoordinatized—without any violation to their topological structure—in such a way as to instantiate a solution to equations of other covariance groups. In the case of complex-valued planewave solutions, the Schrödinger equation played the role of a rival systematization; in the case of real-valued planewave solutions, the Squaringer, free wave, and Schmein–Gordon equations played that role. However, the main lesson of this chapter has been that one needn't venture far into the more complicated solutions of the Klein–Gordon equation in order to find candidate examples of $\langle P, \mathcal{T} \rangle$ or $\langle P_{\mathbb{C}}, \mathcal{T} \rangle$ for which rival systematizations are not so easily found. For example, this was shown to be the case for linear combinations of a small number of real-valued planewave solutions. In particular, the uniqueness of any given equation's dispersion relation (between wavenumber and frequency) entails a uniqueness in the angle of

intersection of overlapping crests for a linear combination of planewave solutions to that equation. For a sufficient number of planewave summands for the Klein–Gordon equation, each having distinct wavenumbers, this would seem to eliminate almost any possibility of finding another equation for which a linear combination of its own planewave solutions could be homeomorphic to it.

It was also shown that for solutions to the general Klein–Gordon equation, even the simplest potential terms brought about changes that make finding a rival systematization more difficult. In certain cases, a rival systematization could be found when the value for m_0 was changed; in others, when the values of the potential terms themselves were changed. But to take advantage of this strategy requires that mass, or the potentials themselves, be taken to supervene. Thus finding rival systematizations would seem to involve an even bolder supervenience claim. Finally, it was shown that some solutions to the general Klein–Gordon equation include values outside the range of $-A \leq \psi \leq A$. And while similar effects could be brought about in solutions to equations with other covariance groups, it would be much more difficult to ensure that the end result of those effects still comprised the same set of field values in the first place. And of course it should not be overlooked that taking a linear combination of such solutions to the general Klein–Gordon equation would bring about even further obstacles, along the lines above.

All of this would suggest that the constructivist's best-systems claim is more plausible than it might have seemed before considering some specific examples. What's more, it has been shown to be plausible in the relatively disadvantageous context of taking Euclidean topological structure as primitive. The enormous symmetry group of $\mathcal{R}^4$ allows a great number of alternative coordinatizations ϕ' to be considered, making it infeasible to argue that all possible rival systematizations have been considered. If a richer topological structure were taken as primitive, and if its symmetry group could be outlined in terms of specific coordinate transformations, then it would be much easier for the constructivist to ensure that all rival sysetmatizations had in fact been considered. Along those lines, one potentially useful topological structure is mentioned in the following chapter. And as

an even stronger example, one might consider the constructivist project in the context of some richer and more familiar fundamental spacetime structure, such as the standard affine structure of $\mathbb{R}^4$. Choosing this as the primitive geometrical structure for one's Humean mosaic would require its own motivation, since the preceding chapters have emphasized the fittingness of taking some topological structure as primitive. But one thing is clear: the affine structure of $\mathbb{R}^4$ is a sub-metrical structure that would in fact leave the metrical structure of Minkowski spacetime to be accounted for by the symmetries of the 'liberalized' best-systems dynamical laws. And in this context, the constructivist project is on very sure footing. The details of this argument are given in Appendix C, where it is shown that if a linear combination of just a few real-valued, planewave solutions to the Klein–Gordon equation are ascribed their 'induced' affine structure of $\mathbb{R}^4$, the resulting space of scalar-field values is arguably best-systematized by a Lorentz-covariant equation—namely, the Klein–Gordon equation itself.

6 Review, Commentary, and Open Questions: On the Constructivist Project of the Dynamical Approach

It was claimed in the Introduction that this thesis would investigate the interplay between explanatory issues in special relativity and the metaphysical foundations of its various interpretations. Along the way, certain metaphysical issues have been addressed, but the culmination of those discussions has been a more mathematical project—namely, exploring the plausibility of the claim that a collection of field values, when ascribed some naturally motivated sub-metrical structure, could support a best-systems account of Minkowski geometry. In this chapter, the metaphysical details of that project are revisited in more detail.

In §6.1, some clarifications are made regarding the constructivist's fundamental ontology and ideology, particularly the commitment to independently existing physical fields, the denial of substantival spacetime, and the primitive relations involved in taking topological structure as primitive. Here also, some remarks are given regarding typical concerns with a 'best systems' explanation involving claims about possible coordinatizations. In §6.2, it is asked whether a regularity relationalist understanding of the dynamical approach could be applied beyond the context of scalar fields and special relativity—for example, in the context of vector or tensor fields, or the variably curved spacetime structure of general relativity. Finally, in §6.3, a concluding summary of the preceding chapters is given, which is supplemented with pointers to various open questions and avenues for future research.

6.1 Metaphysics of the Dynamical Approach

As explained in Chapter 3, the dynamical approach to relativity involves certain explanatory claims that give it a unique position in the spectrum of traditional approaches to spacetime geometry. In Chapter 4, it was argued that those explanatory claims, and the unique approach to spacetime geometry that they entail, are made manifest in a relativistic version of Nick Huggett's 'regularity relationalism'. The basic metaphysical

picture associated with this approach to spacetime structure was outlined in §4.2, but certain issues merit further discussion.

Field Values and Ideological Relationalism

At the close of §2.3, before any introduction to regularity relationalism, it was argued that the dynamical approach to relativity is most naturally understood as a form of ontological relationalism, meaning that spacetime is not taken to exist—or at the very least least, that it is not taken to exist independently of matter. It was also mentioned in §4.3 that this would mean the physical fields themselves ought to be understood to exist independently of spacetime.[1] To clarify the constructivist's conception of fields, it can be contrasted with two conceptions of fields available to spacetime substantivalists.[2] One of these is summarized well by Hartry Field.

> What is a field theory? As I see it, a field theory is simply a theory that assigns causal properties to space-time points or other space-time regions directly (as opposed to indirectly, via matter that occupies these points or regions). [...] Obviously this presupposes a substantival view: on a relationalist view, there are no points or other regions of unoccupied space-time, so the assignment of a property to such a point or region makes no sense.
>
> [Field 1984, p. 40; *cf.* 1980, p. 35]

Field's view is synonymous with what Earman calls 'supersubstantivalism' [1989, p. 115]. But rather than assign field values to points of spacetime, the dynamical approach includes the point-like parts of physical fields themselves in the fundamental ontology. More will be said below about the distinction between supersubstantivalism and the dynamical approach construed as a relativistic version of regularity relationalism. But first, another spacetime substantivalist's conception of fields is worth distinguishing—namely, the one according to which physical fields reside *in spacetime*. In contradistinction, the point-like parts of physical fields under consideration in §§4.2–4.3 and in all of Chapter 5 are not to be understood as being located somewhere in spacetime, or even as coexisting with

[1] See footnote 24 on page 115.
[2] In recent and forthcoming work, Pooley makes this distinction [2013, §6.1.2; 2014, §5.8].

substantival spacetime in any way. Rather, they are to be understood as the sole elements of the fundamental ontology.[3]

That being said, it should be acknowledged that the explanatory claims of the dynamical approach do not necessarily commit its advocates to this conception of fields, or to ontological relationalism at all. Rather, those explanatory claims merely commit the constructivist to what was called 'ideological relationalism' at the close of Chapter 2. That is to say, following on from footnote 31 on page 65, the dynamical approach would in fact be compatible with a particular form of substantivalism, according to which the fundamental ontology includes a substantival manifold with sub-metrical structure, from which the inhabiting physical fields inherit their own geometrical structure.[4] The best-systems argument could be carried out as before, so that Minkowski geometry supervenes upon the symmetries of the (liberalized) best-systems dynamical laws. And while in this case the constructivist project would certainly not give an account of the primitive geometrical structure of spacetime *the entity*, it would still give an account for the geometrical structure that best codifies the behavior of material objects. That is, the constructivist project would give an account for the supervenient geometrical structure that substantival spacetime has in virtue of the way that material objects behave in it.[5] But still, the in-principle compatibility of the dynamical approach with this unusual form of substantivalism should not distract us from the construal of the dynamical approach presented in the preceding chapters. The idea that physical fields, existing in a sub-metrically structured substantival spacetime, could behave in accord with dynamical laws whose symmetries subvene a richer geometrical structure for that substantival spacetime, is not well motivated. The constructivist would certainly be wise to keep the supervenient geometrical structure associated with the same entities whose

[3]Incidentally, this kind of view is also acknowledged by Field [1984, p. 41].

[4]Because this form of substantivalism would not take the metric itself to represent spacetime, it would be a form of so-called 'manifold substantivalism'. An introduction to the metaphysical details of this and other conceptions of substantivalism is provided by Hoefer [1996, §I].

[5]Most of Brown's claims about geometrical structure codifying the behavior of material objects would appear to be open to interpretation along these lines [2005*b*, p. 9, p. 100, p. 133, p. 142], although it should be stressed that there is no reason to suggest that Brown does in fact have anything like this in mind.

dynamical laws have the symmetries that account for it, and this brings one back to a relationalist construal of the dynamical approach.

However, it should also be acknowledged that the field-points under consideration in §§4.2–4.3 and in all of Chapter 5, which amount to nothing more than the point-like instantiation of some field value, seem to have no more than a mere terminological difference from the points of spacetime in the aforementioned *super*substantivalist's perspective. For in either case, the fundamental ontology and ideology amounts to a set of point-like entities with primitive scalar-valued properties.[6] Thus an argument might be made to understand the dynamical approach as a form of supersubstantivalism, according to which the geometrical structure of spacetime supervenes upon the symmetries of the 'liberalized' best-system dynamical laws for the degrees of freedom of the spacetime points themselves. And yet, while the difference between the relationalist approach described in the preceding chapters and the supersubstantivalist one being suggested here may boil down to a terminological difference in the case of a single scalar field, this is not the case when dealing with multiple physical fields. For the supersubstantivalist, a system comprising multiple physical fields is not different, ontologically speaking, from a system involving a single field. The only change that comes about in going from one to several fields is in the ideology—the number of fields determines the number of scalar-valued properties instantiated by each point of spacetime. But what the supersubstantivalist still calls a point of spacetime, the dynamical approach must now describe as a set of field-points standing in some coincidence relation.[7] And while these two entities—the supersubstantivalist's point of spacetime and the relationalist's coincidence of field-points—might have degrees of freedom that are given the same description mathematically, there is certainly metaphysical difference between them. At the very least, there is a difference in number when comparing a single entity having n degrees of freedom with a collection of n entities that each have a single degree of freedom, and each of which stands in some

[6] The exception would be if the relationalist did not assume fields to exist in regions in which their values are 0. Pooley discusses related issues [2013, pp. 573–4].

[7] In Appendix B, it is shown that this coincidence relation can actually be reduced to topological indiscernibility.

coincidence relation with each other.

Fundamental Topological Structure

As explained at the close of Chapter 2, and as mentioned when considering alternative topologies in §4.2, the constructivist's fundamental field-points cannot primitively instantiate the geometrical structure that is to be accounted for by the symmetries of the dynamical laws. Instead, the collection of fundamental field points must be metrically amorphous. Only some sub-metrical structure can be taken as primitive, and it was suggested above that a topological structure should be considered. However, taking topology as primitive raises some metaphysical questions, which deserve more attention.

Before continuing, one clarification should be made: the constructivist's reason for taking topological structure as primitive is not simply to minimize the fundamental ideology of the dynamical approach. In other words, it is not as if topological structure is a sufficiently rich structure for our best physics, and that it is always, tacitly assumed whenever metrical structure is taken as primitive, and that the constructivist therefore wishes to excise metrical structure from the fundamental ideology by some application of Ockham's razor. Granted, on a mathematical level, topological structure can be recovered from any metrical structure, and so it would seem that topology comes as part of the package in any metrical structure. But on a metaphysical level, it is not as if primitive topological properties or relations are tacitly ascribed to material objects whenever metrical structure is taken as primitive. Cian Dorr makes a similar point in a response to recent work by Tim Maudlin, which will be introduced further below.

> It seems to me that the mathematical sense in which topology is said to be 'more fundamental' than affine and metric geometry is quite different from the metaphysical sense of 'fundamental' we are concerned with. The mathematical 'fundamentality' of topology is a kind of generality: there are many kinds of mathematical structure within which there are natural definitions of properties obeying the topological axioms for 'open set'; this makes topology useful for capturing behaviour common to many mathematical structures. There is nothing in this to count against the hypothesis that the metaphysically

> fundamental facts about physical space are all facts about its affine or metric structure. [Dorr 2011, pp. 144–5]

Dorr's point seems to be that, when it comes to choosing some spatiotemporal properties or relations to be included in one's ideology, there is nothing to suggest that topological ones are any more simple or basic than metrical ones. And on this point the constructivist can agree: the reason for taking topology as primitive is not simply to minimize or simplify the spatiotemporal component of one's ideology. Rather, what motivates the constructivist's choice of taking topology as primitive is threefold: first, it is arguably a natural and fitting choice for Lewis's modal realism that is so closely related to Humean supervenience; second, topological structure is more naturally motivated than metrical structure;[8] and perhaps most importantly, it is a sub-metrical structure from which a best-systems account of Minkowski geometry can be given, so that the arrow of explanation runs from dynamics to geometry. However, while metrical structure can be captured by relatively simple, two-place 'distance' or 'interval' relations on field-points, topology has at its core the somewhat unintuitive concept of an 'open set'. So while the mathematical details of taking topological structure as primitive have been addressed in Chapter 5 and Appendix A, the metaphysics of taking topology as primitive deserves a bit more attention.

First of all, given that the constructivist project has been introduced in terms of a fundamental ontology and ideology, and given that Quine defines the words 'ontology' and 'ideology' in terms of variables bound by first-order existential quantifiers, it might be asked whether topological structure can in fact be cashed out in terms of first-order relations.[9] Indeed, in what kind of relations do the field-points of a topologically structured space stand? Given the centrality of open sets in topology, it would seem that they must stand in some kind of 'open set' relations with each other. But unlike the binary 'interval' relations ascribed to the points of a metrically structured space, some such 'open set'

[8] See the discussion in §3.2; Schlick's 'method of coincidences' is illustrative of this fact.

[9] Quine's terminology is introduced in footnote 2 on page 1, above. The label 'first-order' is used to refer to those quantifiers that bind variables standing for objects, not properties. That is, first-order quantifiers quantify elements of the ontology, and not to elements of the ideology.

relation would have to hold between collections of numerous field-points. To be clear, it would be a mistake to think that the members of an open set each stand in some 'open set' binary relation with each other member of the set. A more accurate description of taking topology as primitive would be to say that the members of an open set collectively bear an 'open set' property. But whereas first-ordered logic allows for quantification over individual objects, this notion of a collectively borne 'open set' property involves quantification over groups or sets of the fundamental ontology. Situations like this are the motivation for adopting plural quantifiers into first-order logic, which bind multiple variables that take the values of objects, but not properties [Linnebo 2013]. Thus in the standard presentation of topology, for field-points to instantiate some topological structure would require a notion of what is called plural quantification. And with that one caveat, taking topological structure as primitive can indeed be cashed out in terms of first-ordered logic, since the plural 'open set' property is held by (sets of) elements of the ontology, and not of the ideology.[10]

At this point, however, it might be thought that another problem arises in attempting to cash out topological structure in terms of first-order quantifiers. For in the standard presentation of topology, sets must satisfy certain axioms in order to instantiate a topological space, and these axioms involve ways that the sets relate to each other.[11] But if the constructivist's fundamental ontology includes only field-points, then to refer to relations between open sets is to admit relations between elements of the ideology—i.e. relations that are held by not the objects themselves but by their properties. And so even if a pluralist 'open set' relation is included in the constructivist's ideology, it would seem that some further, second-order relations must be involved in taking topology as primitive. But in fact this isn't the case. For while the axioms of a topological space do indeed involve statements about the relationship of open sets with one another, this does

[10] It should be noted that there is nothing about plural quantification that goes against Humean supervenience, which lies at the heart of the dynamical approach as described above. What bears the property of 'open set' is a plurality of the point-like members of the fundamental ontology. Indeed, Lewis himself advocates plural quantification in his work on set theory [1993, pp. 9–12]. Outlines of Lewis's work on set theory are provided by Weatherson [2010, §7.1] or Burgess [2014].

[11] See Appendix A for details.

not mean that the open sets are required to actually stand in some particular relation on a metaphysical level. Rather, the axioms of a topological space have to do with how the points of the space instantiate the 'open set' relation, and what characteristics of the open sets will ensure that those 'open set' relations are being held in such a way that the points do indeed have some topological structure. So while describing the characteristics of a topological space, or the axioms that must be satisfied by the open sets of a topological space, might require second-order language, this does not entail any second-order elements of the constructivist's fundamental ideology.

Laws and Explanation in the Dynamical Approach

In Chapters 3 and 4, it was the explanatory claims of the dynamical approach that led to the suggestion that it be interpreted as a form of regularity relationalism. According to this construal of the dynamical approach, Minkowski geometry supervenes upon the symmetries of the liberalized best-system dynamical laws. But whether this form of explanation suffers any serious problems on its own should also be considered. For instance, one of the most common concerns with a best-systems account of any structural feature of the world is that it seems to amount to the following claim: while the world may appear to have some structural property, it simply does not; rather, it is merely possible to coordinatize the actual world in the same way that one would if it did have that structure, and to enjoy the benefits of using dynamical laws written in terms of that structure. In particular, a best-systems account of Minkowski geometry seems to acknowledge all of the evidence for that structure being a primitive part of the world, while stubbornly refusing to admit it. Instead, all that is granted is the existence of some coordinatization of the material world, with respect to which dynamical laws can be written which would suggest the existence of that structure.

In introducing regularity relationalism above, it was mentioned that less favorable responses have come from others, including Cian Dorr. And in fact, it is this very aspect of a best-systems explanation that he criticizes. In a larger work on nominalism, Dorr argues

against theories that use existential quantifiers to eliminate primitive commitments, and points out that this seems to be precisely what is going on in regularity relationalism.

> There is a general pattern here. If we want to weaken a theory so as to eliminate its commitment to some sort of hidden structure, we can often do so by replacing the vocabulary which purports to characterize this structure with variables of an appropriate sort bound by initial existential quantifiers. Philosophers who are suspicious of particular putative bits of hidden structure keep on rediscovering this fact, and announcing that they have shown how to eliminate the structure in question. But once we have realized the complete generality of the trick, we should not be impressed by their achievements.
>
> [Dorr 2010, p. 160]

As a particular example, Dorr picks on Huggett's regularity relationalism for accounting for absolute acceleration by adopting "a theory of the form, 'There are some admissible coordinate systems in which Newton's laws hold', where 'admissible' coordinate systems are those that respect the history of inter-particle distances" [p. 161]. And, perhaps even more importantly for our purposes, Dorr also picks on the idea of reducing geometry to topology.

> Ordinary physical theories formulated in coordinate terms are naturally understood as claiming that space-time has a geometric structure much richer than that of mere topology: regions of space-time can differ in all sorts of intrinsic geometric respects even when they are topologically indiscernible. But any such theory can be weakened so as to remove this implication: we need only say that *there is some coordinate system* which respects the space-time's topological structure, relative to which the given dynamical equations are true. This minimalistic way of thinking about the content of these theories is favoured by van Fraassen [1970].[12] And, under the influence of a positivist philosophy that rejects the whole idea of intrinsic structure, similar ideas are still sometimes found in physics textbooks—it is common to formulate Newton's first law as the claim that 'There exist inertial frames', where these end up getting defined as frames in which Newton's laws hold (see, for example, Woodhouse [2003]). But I have no interest in deferring to the opinions of physicists when these are manifestly influenced by the discredited anti-realist philosophies of the past. We do have good reason to ascribe to

[12] See the discussion on page 98 regarding van Fraassen's formulation of Newton's first law.

> space-time an intrinsic geometric structure that goes far beyond topology. Thus the existentially quantified theories which purport to explain all our observations without entailing that there is any such structure must be bad theories. [Dorr 2010, p. 161]

In reply, it should first be pointed out that whatever van Fraassen's motivations may have been, positivistic thinking is not at the heart of the proposal to interpret the dynamical approach as a form of regularity relationalism. It has already been shown in Chapter 3 that Brown's work does not in fact fit in the positivistic approach to spacetime structure that Einstein is sometimes taken to support by presenting his 'practical geometry' against Poincaré's conventionalism. What's more, whereas Dorr's target would "reject the whole idea of intrinsic structure", the constructivist grants intrinsic topological or differentiable structure. And whereas Dorr's target would also deny the metrical structure of spacetime, we must keep in mind that Brown does in fact acknowledge it, but merely wishes to give an account for it. Recall, for example, the sort of statements outlined in Chapter 1 above, like "The appropriate structure is Minkowski geometry *precisely because* the laws of physics of the non-gravitational interactions are Lorentz covariant" [2005*b*, p. 133; *cf.* Brown and Pooley 2006, p. 80]. As mentioned above, it was this explanatory claim that led to the suggestion that the dynamical approach be interpreted as a form of Huggett's regularity relationalism. And Huggett, for his part, has also made it clear that his motivations are not positivistic. Rather, they are "the usual ones: worries about 'Leibniz shifts' and considerations of ontological parsimony that mitigate against the introduction of bizarre non-material substances" [2006, p. 41] And so even if the strategy of weakening a theory's metaphysical commitments by the use of existential quantifiers has its roots in a branch of philosophy no longer in trend, we mustn't throw the baby out with the bathwater; the motivation here is different from that of Dorr's positivist target.

Second, we might enquire as to what 'good reasons' Dorr has in mind for ascribing to spacetime (or to the relationalist's material ontology) an intrinsic geometric structure that goes far beyond topology. Presumably, one of those reasons would be the fact that certain phenomena like length contraction could be explained by postulating some such

structure. But of course whether intrinsic geometrical structure can in fact play some such explanatory role is central to Brown's thesis—that is precisely the idea that he challenges. Another 'good reason' might simply be that our best dynamical laws are written in terms of such structure. But again, the lesson of Huggett's work on regularity relationalism is that with a 'liberalized' best-systems conception of laws, one needn't be committed to the full spatiotemporal structure in terms of which the dynamical laws are written.

In closing, it should be noted that Dorr makes a similar criticism of what he calls 'as-if' laws elsewhere [2011, pp. 146–7], and Frank Arntzenius gives a similar critique of best-systems explanations à la Huggett, citing Dorr's work quoted above [Arntzenius 2012, pp. 169–70]. The gist of these argument is roughly the same, and can be conveyed through the example that Arntzenius gives.

> My only objection to such a theory [regularity relationalism] is that such a theory is not a simple theory. By 'simple' I mean in the sense that is important when it comes to evaluating whether a theory is likely to be true, or likely to be approximately true. [...] This argument is that if such theories were to count as simple it would make life too easy for people who want to get rid of objects and properties that they do not like for some whimsical reason. For instance, suppose I were to claim that the world is pretty much as you think it is, except that rocks do not exist. I then go on to claim that the true theory of the world is that there are no rocks, and that the true theory of the world merely says that the things and properties and relations that there are [...] are embeddable in a non-existing make-belief world which includes rocks and in which your favorite make-belief laws hold. It seems obvious to me that such a theory is not simple in the sense relevant to evaluation [of] how good such a theory is, and that one should not believe that it is true. Theft is theft. Honest toil is honest toil. [Arntzenius 2012, pp. 169–70]

And here, a point mentioned briefly above bears repeating. The target of Arntzenius's criticism would seem to make the claim that there is 'some coordinatization' with respect to which the field-points stand in the same spatiotemporal interval relations that they would if they were embedded in Minkowski spacetime, and would use this as a way of *excising* Minkowski spacetime structure from the theory. But by taking a (liberalized)

Humean account of dynamical laws, the dynamical approach does not so much purge Minkowski geometry from the picture as it does change it from a primitive structural feature to a supervenient one. Under this approach, there is a fact of the matter about what the best-systems dynamical laws are for the Humean mosaic, and Minkowski geometry (along with many other things) supervenes upon those laws being Lorentz-covariant. What's more, this project is not nearly as simple as Arntzenius makes it out to be. As should be clear from the project undertaken in Chapter 5, it is not enough to show that there is 'some coordinatization' according to which Lorentz-covariant laws hold; rather, it must also be argued that those are the (liberalized) best-system dynamical laws.[13]

6.2 Going Further with the Dynamical Approach

In Chapters 4 and 5, the systems under consideration were admittedly quite simple: single scalar fields with Euclidean topological structure. And of course these are not realistic examples; worlds like ours are much more complex, comprising multiple fields with vector or tensor components and exhibiting a variably curved geometrical structure. Granted, the purpose of the case study carried out in Chapter 5 was not to come up with a mathematical model of our own world, but only to defend the plausibility of the constructivist project. But still, if the dynamical approach is to be taken seriously as an interpretation of relativity, it should be asked whether its tactics are applicable in more general contexts.

With regard to working with different types of fields, this would amount to a change in the constructivist's fundamental ontology. On the one hand, taking multiple fields as primitive poses no real problems for the constructivist, although as mentioned above, some form of a coincidence relation might have to be added to the ideology.[14] But as for taking vector or tensor fields as primitive, some more serious obstacles will need to be overcome. Those are addressed below. Finally, separate and aside from the type of

[13] Regarding Arntzenius's final two sentences, see also Huggett's own comments regarding the objection that regularity relationalism amounts to 'theft' [2006, p. 46].

[14] As mentioned above, a strategy for the constructivist project is outlined in Appendix B, according to which coincidence can be made redundant with the right choice of topology.

fields under consideration, it is asked how well the dynamical approach *qua* a relativistic form of regularity relationalism could fare in accounting for variably curved geometrical structures.

Vectors and Tensors

In order to consider supporting a liberalized, best-systems account of Minkowski geometry with primitive vector or tensor fields, one might consider enriching the ontology and ideology straightforwardly. Consider, for example, the case of a single vector field. Instead of assuming the existence of a set P of scalar field-points, one would assume the existence of a set $P_{\uparrow}$ of vector field-points. As vectors are typically construed as geometrical objects on a differentiable manifold, the set $P_{\uparrow}$ should be assigned some differentiable structure, rather than mere topological structure.[15]

Before going on, a few words are in order regarding the standard characterization of differentiable structure. In moving from topological to differentiable structure, what changes to the constructivist's ideology are being assumed? As discussed above, topological structure is characterized in terms of open sets, which in turn were accounted for above in terms of a primitive, pluralist property. As shown in Appendix A, differentiable structure is characterized by an 'atlas' $\mathcal{D}$ of coordinatizations from the space into $\mathbb{R}^4$, with those coordinatizations satisfying certain differentiability requirements. Thus the topological space of scalar field-points $\langle P, \mathcal{T} \rangle$ in Chapters 4 and 5 would be replaced by $\langle P_{\uparrow}, \mathcal{D} \rangle$. But as for what other properties or relations are required in cashing out a primitive differentiable structure on the field-points, it is not clear that anything is needed beyond the aforementioned 'open set' property. That is to say, differentiable structure can effectively be understood as a particular form of topological structure, so that taking differentiable structure as primitive is comparable to taking any particular topological structure as primitive. Consider, for example, the working assumption in

[15] As mentioned in the preceding chapters, taking differentiable structure as primitive is just as fitting for the constructivist project as is taking topological structure as primitive. Like topological structure, it is a naturally motivated and highly symmetric sub-metrical structure that leaves plenty of geometrical structure to be accounted for by the symmetries of the dynamical laws.

Chapter 5 that $\mathcal{T}$ was a Euclidean topology. This assumption amounted to there being open sets which not only satisfied the basic axioms of a topological space, but which did so in a way that enabled a global homeomorphism $\phi : \langle P, \mathcal{T} \rangle \to R$, with R being some open subset of $\mathbb{R}^4$ with the standard Euclidean topology $\mathcal{R}^4$. In other words, for $\mathcal{T}$ to be a Euclidean topology is for the open sets to be instantiated by the field-points in such a way that this mapping is possible. Similarly, in taking a differentiable structure as primitive, the constructivist assumes that the open sets satisfy the basic axioms of a topological space in such a way that, regardless of whether there is a global homeomorphism into a subset of $\mathbb{R}^4$, there is a local homeomorphism into some such subset for at least one neighborhood of every field-point, and those local homeomorphisms will satisfy the given differentiability requirements.[16] In both cases, the only spatiotemporal component of the constructivist's fundamental ideology is a pluralist 'open set' property. How those open sets are instantiated will determine what kind of sub-metrical structure is taken as primitive.[17]

But what of the question of taking vector fields as primitive? Almost immediately, the project would seem to face problems. For unlike scalar field-points, these vector field-points would have degrees of freedom that could not be described by a simple mapping into $\mathbb{R}$. And however one chooses to illustrate their degrees of freedom, a circularity arises: the definition of the vector field-points p, which are to be assigned some differentiable structure, mentions that structure itself. For instance, one common description of vector field values is the mapping $v_f : \mathbf{C}_f(p) \to \mathbb{R}$ on the collection of all possible smooth, real-valued functions f that pass through p. In order to codify the degrees of freedom for a vector field, this mapping must satisfy certain conditions, which needn't concern

[16] Those requirements are outlined in more detail in Appendix A.

[17] It should be noted that Arntzenius and Dorr outline an alternative characterization of differentiable structure, which does not involve local homeomorphisms into $\mathbb{R}^4$ [Arntzenius 2012, §8.4 *cf.* Penrose and Rindler 1984, §4.1]. While interesting in its own right, and clearly of potential use for the authors' nominalist project, this characterization would entail a significant change in the constructivist's fundamental ontology, since it does away with open sets as primitive terms. The rest of their coauthored chapter offers other interesting alternatives, including a three-place predicate that would determine a space's differentiable structure [2012, §8.6, §B]. These too involve substantial changes to the ontology that are motivated by the authors' interest in nominalism.

us here.[18] The salient point is that v_f acts upon *smooth* functions passing through p, meaning that this description of a vector field's degrees of freedom assumes some kind of smoothness structure from the start. As another example, we might consider describing vector field-points' degrees of freedom by a mapping $v_t : P_\uparrow \to \cup_i T_i$, which maps each field-point p to an element of the tangent space T_p located at that point. But because tangent spaces are defined in terms of the differentiable structure of a manifold, here again the description of the field-points' degrees of freedom tacitly assumes some differentiable structure. In short, the problem is simply that the constructivist does not seem to be able to say, without referring to the the field-points' collective differentiable structure, *what things* are to be assigned some differentiable structure. Pooley has raised this concern in another context.

> More complex fields [than scalar fields] can pose additional problems for the relationalist. Standard vector and tensor fields, for example, are not obviously conceptually independent of the structure of the manifold on which they are defined. Their degrees of freedom at a point are normally understood as taking values in the *tangent space* at that point (or in more complex spaces constructed in terms of it), which might appear to presuppose the differentiable structure of the manifold on which the fields are defined. [... There] is something suspiciously circular about taking the spaces in terms of which a field's degrees of freedom are defined to be themselves defined in terms of that field's own spatiotemporal extension. [Pooley 2013, pp. 555–6]

Now, if Pooley's point is granted, this would leave the constructivist scrambling for some way to make sense of taking a vector or tensor field as primitive, so that its degrees of freedom could be understood without recourse to the differentiable structure being ascribed to them. One possibility would be for the constructivist to make an even stronger supervenience claim, so that the very aspects of a vector or tensor that require differentiable structure to describe, could in fact supervene upon some best systematization of degrees of freedom that do not. For example, the constructivist might let the 'directed' aspects of a vector field-point, the definitions of which involve reference

[18] Examples are outlined by Malament [2012, p. 8]. For instance, v_f may be the directional derivative.

to differentiable structure, be given a best-systems account in terms of point-like entities whose degrees of freedom are scalar-valued. The constructivist would admit only the vectors' four-scalars (magnitudes) as primitive, and give a best-systems account of the four-scalars' 'components' in any given coordinatization. But as interesting as such an approach might be, any move along these lines would certainly come with the cost of making the best-systematization claim even stronger. In §5.3, it was acknowledged that moves like this, which make more of the system under consideration supervene upon the best-systems dynamical laws, make the constructivist project appear less plausible.

Fortunately, it is not clear that Pooley's circularity concern should prevent the constructivist from taking vector or tensor fields as primitive. Consider, for example, why it is that the same concern would not apply to a spacetime substantivalist: the differentiable structure could be assigned to the point-like parts of spacetime itself, and then the field's degrees of freedom can be defined in terms of it. But it should also be recalled from above (p. 158) that in the context of a single field, the regularity relationalist construal of the dynamical approach differs only terminologically from a version of manifold supersubstantivalism. And if that's the case, one might wonder why it is that only the former runs into trouble when assuming the existence of vector or tensor fields. The problem, it would seem, comes up in attempting to describe the Humean mosaic. For unlike in the case of substantivalism, the constructivist has no way of describing the entities to which some differentiable structure is being assigned without tacitly referring to that structure 'before' assigning it. But really the trouble lies in the word 'before'. Granted, it may not be possible to *describe* the constructivist's vector field-points without reference to their collective differentiable structure, but it is not clear that the constructivist has any need to do so. The constructivist does not begin with a set of vector field-points and *then* assign them some differentiable structure to them; rather, what is taken as primitive is a differentiably structured set of point-like entities whose degrees of freedom can only be described in terms of that differentiable structure. While this might open the door to circular descriptions of the constructivist's

ontology, this does not commit the constructivist to believing in the possibility of vector field-points *sans* differentiable structure. The only difference between the constructivist and the supersubstantivalist is the former's commitment to the fact that if the vector field were removed from the picture, no differentiably structured entity would remain. And so as long as the constructivist is not misunderstood as being committed to the existence of $P_\uparrow$ alone, but rather to the existence of $\langle P_\uparrow, \mathcal{D} \rangle$, there should be no concern with enriching the constructivist's fundamental ontology with vector or tensor fields.

The Dynamical Approach to General Relativity

The chapters above have been written primarily in the context of special relativity. And so they should be: Brown and Pooley's dynamical approach to relativity does the same, as does most of the literature outlined in Chapter 1. But of course it should also be asked how to understand the dynamical approach in the context of general relativity.

In fact, this question was raised as early as 2001, in Brown and Pooley's promotion of the 'truncated Lorentzian pedagogy'. There, the authors began by outlining a relatively straightforward extension of the Lorentzian pedagogy to general relativity, which amounted to the claim that even there, local geometrical structure might be accounted for by the local validity of special relativity. The idea could be summed up in their maxim "no matter, no metric", meaning that the vacuum solutions to Einstein's field equations would be taken to have a physically meaningful affine structure, but no physically meaningful metrical structure [p. 269].[19] But as Brown and Pooley acknowledge, this approach would have serious problems.

> It overlooks the simple fact that the metric field in GR [...] appears to be a *bona fide* dynamical player, on par with, say, the electrodynamical field. Even if one accepts [...] the 4-metric as less fundamental than the evolving curved 3-metrics of the Hamiltonian approach to GR, it is nonetheless surely coherent to attribute a metric field to spacetime whether the latter boasts matter fields or not. [Brown and Pooley 2001, p. 270]

[19] See also earlier work by Brown on this topic [1997].

And so unlike in the case of special relativity, the authors would grant that in general relativity, the metric field exists independently of any matter fields, which would preclude any attempt to make it supervene upon features of the dynamical laws.

Despite letting go of the arrow of explanation from dynamics to geometry in a generally relativistic context, Brown and Pooley go on to argue that the Lorentzian pedagogy is still the right approach to giving an account for how rods and clocks survey metrical structure. For while a generally relativistic metric field may define a locally Minkowskian geometry on small regions, the authors take it to be a *non sequitur* that the latter should give an account for the Lorentz covariance of non-gravitational dynamical laws [p. 270]. What the authors take to account for relativistic phenomena like length contraction and time dilation is what might be described as the generally relativistic version of Einstein's 'restrictive principle'.[20]

> The assumption, which is intimately related to the postulate of minimal coupling in GR, is that relative to the local Lorentz frames, insofar as the effects of curvature (tidal forces) can be ignored, the laws for the non-gravitational interactions take their familiar special relativistic form: in particular, the laws are Lorentz covariant. [Brown and Pooley 2001, p. 271]

However, letting go of the arrow of explanation from dynamics to geometry in a generally relativistic context does not mean that Brown would go so far as to admit that generally relativistic systems are embedded in substantival spacetime with some primitive geometrical structure. For while it may no longer be argued that geometrical structure is accounted for by the symmetries of the dynamical laws, it could still be argued that the metric ought not be understood as a structure of spacetime itself, but rather as a physical field that receives its geometrical characteristics by the way that it couples with matter fields. This approach appears primarily in Brown's later work, where special emphasis is put on this distinction.

> [In] GR the metric tensor (and hence the affine connection, given their compatibility) becomes a dynamical agent. It undeniably acts and is acted

[20]The 'restrictive principle' associated with Einstein's 1905 principle-theory is introduced in §1.1.

> upon. But in what sense is it *geometry*, or *space-time structure* itself that acts? [Brown 2005*b*, p. 150]

Indeed, the final chapter of Brown's book is framed as an attempt to develop a less familiar understanding of metrical structure in general relativity, outlined originally by Arthur Eddington [1923, §64], according to which spatiotemporal intervals are understood in terms of measurements made with material objects [Brown 2005*b*, pp. 150–1]. Brown's aim in doing so is to show that not only does the dynamical approach still hold in general relativity when it comes to kinematical phenomena (as described above), but also that while the metric field might be granted independent existence, its geometrical 'significance' hinges upon the way that matter couples to it.[21]

> It is because of minimal coupling and local Lorentz covariance that rods and clocks, built out of the matter fields which display that symmetry, behave as if they were reading aspects of the metric field and in so doing confer on this field a geometric meaning. [Brown 2005*b*, p. 176]

Thus the arrow of explanation from dynamics to geometry takes a very different shape in the context of general relativity.[22] Rather than saying that geometrical structure supervenes upon the symmetries of the dynamical laws, Brown would argue that the metric field is only interpreted as geometrical structure due to the way that material objects interact with it.

Now, it was mentioned above that at least in the context of a single physical field, the regularity relationalist construal of the dynamical approach to special relativity would appear to differ only terminologically from a version of supersubstantivalism. And in the context of general relativity, another comparison can be made between the dynamical approach and substantantivalism. For as Pooley has noted, it is not clear how the dynamical approach to general relativity, described above, actually differs from a traditional substantivalist view.

[21] See also Skow's comments on page 75, above.

[22] As Pooley notes [2013, §7], this is a common move for 'have-it-all' relationalism in the transition from special to general relativity.

> Both parties accept the existence of a substantival entity, whose structural properties are characterized mathematically by a pseudo-Riemannian metric field and whose connection to the behavior of material rods and clocks depends on, *inter alia*, the truth of the strong equivalence principle. It is hard to resist the suspicion that this corner of the debate is becoming merely terminological. [Pooley 2013, p. 578]

Pooley goes on to say that one benefit of the substantivalist's position is that it has a certain continuity with its position on the metric in the context of pre-generally relativistic spacetimes, whereas according to the dynamical approach, the status of metrical structure changes significantly in a generally relativistic setting [p. 579]. Not to be outdone in terms of continuity, then, we might ask whether the interpretation of the dynamical approach advocated in Chapters 3 and 4 could be made to work in the context of general relativity, and if so, whether it offers any way to preserve the stronger explanatory claim that geometrical structure supervenes upon dynamical laws' symmetries.

From the start, it should be clear that any such project would face serious difficulties given the variable curvature of generally relativistic spacetimes. That is, in the approach outlined in §4.3, the geometrical structure that supervened upon the best-systems dynamical laws was the one whose symmetry group matched the group of transformations under which the best-systems dynamical laws preserved their form. But in the case of a variably curved geometry, no such symmetry group will arise; it is the absolute, 'static' geometrical structure of a specially relativistic world that allows it to supervene in such a way. It would not be clear how to express the supervenience claim for a set of generally covariant equations that do not take an especially simple form in any coordinatizations related by some specific subgroup—the way that Lorentz-covariant equations do in coordinatizations related by the Poincaré group. And so it seems that in the context of general relativity, we would have to let go of this particular way of having geometry supervene upon the symmetries of the dynamical laws.

But that being said, there is another sense in which a relativistic form of regularity relationalism which takes sub-metrical structure as primitive would seem to find a more

natural home in general relativity than in special relativity. In a related context, DiSalle gives a description of general relativity that makes this clear.[23]

> In the context of general relativity, as understood by Einstein and Schlick, the relativity of geometry has an immediate significance for metrical geometry: the underlying "amorphous" space is represented by an arbitrary Riemannian manifold, assumed to have no more intrinsic structure than its differentiable structure [...]. Physical structure, more precisely the metrical structure that is to play the role of the gravitational field, is imposed by two stipulations: first, that special relativity holds in the infinitely small, i.e. at any point the metric is Minkowskian; second, that over finite regions, the metric depends on the mass distribution in accord with Einstein's equation.
>
> [DiSalle 2006, p. 87]

On this understanding of metrical structure, the constructivist might be tempted to take differentiable structure as primitive and give some kind of best-systems account for a metric that is approximately Minkowskian on a local scale, and globally dependent upon mass distributions. To accomplish this, one might revisit a strategy that failed in Chapter 4—namely, taking as primitive some sub-metrical structure that is richer than the standard topology of $\mathbb{R}^4$, and yet which does not encode the allegedly supervenient metrical structure from the start.[24] For if the metrical structure that will allegedly supervene upon the best-systems dynamical laws is no longer so symmetric in space and time—i.e. if the best-system dynamical laws will not have so simple a symmetry group—then it may be possible to enrich the sub-metrical structure without tacitly taking as primitive the very metrical structure that is meant to supervene.

Consider, for example, the 'Path' topology $\mathcal{P}$, which was developed by Stephen Hawking and collaborators [1976] as an adaption of the Fine topology (Ch. 4) that would be more suitable for general, curved spacetimes.[25] While this is not the place for a full

[23] DiSalle goes too far in stating that the metric at any point is Minkowskian, which would suggest zero curvature at any point. I take DiSalle to mean "approximately Minkowskian" in the following quotation.

[24] Recall that the main obstacle in choosing to take as primitive some alternative, richer topology like the Fine topology $\mathcal{F}$ was that this topology encodes Minkowski metrical structure up to scale, leaving no work to be done by the dynamical laws.

[25] Hawking et. al. and also cite Rüdiger Göbel's work [1976*a*; 1976*b*; 1980] on generalizing Zeeman's results to generally relativistic spacetimes, along the lines of Zeeman's original suggestion [1967].

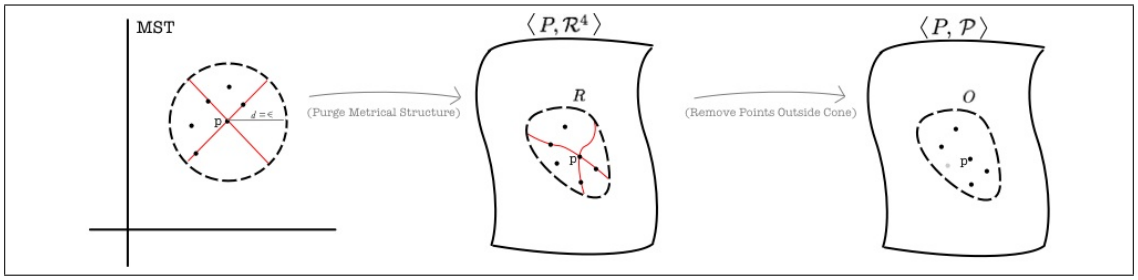


Figure 28: The Euclidean metric d on inertial coordinates of Minkowski spacetime induces $\mathcal{R}^4$. Removing from an open set R the points that lie outside of the lightcone of some point p (shown in red) creates a new set O, which is open in $\langle P, \mathcal{P} \rangle$, but not in $\langle P, \mathcal{R}^4 \rangle$.

introduction to the Path topology, some of its salient features can be described briefly.[26] In short, whereas the Fine topology was defined as the finest topology to induce $\mathcal{R}^1$ on all of its time axes and $\mathcal{R}^3$ on all of its space axes, $\mathcal{P}$ is defined as the finest topology to induce $\mathcal{R}^1$ on any continuous time-like curve. Like $\mathcal{F}$, this means that $\mathcal{P}$ is finer than $\mathcal{R}^4$. And whereas $\mathcal{F}$ was finer in virtue of its having all the open sets of $\mathcal{R}^4$ plus others including those illustrated as in Figure 8 (p. 177), so too $\mathcal{P}$ has all of the open sets of $\mathcal{R}^4$ in addition others including those illustrated in Figure 28. With open sets defined in this way, the Path topology encodes a differential, causal, and conformal structure for spacetime, but no metrical structure. And so whereas the Fine topology encoded the metrical structure of Minkowski spacetime up to scale, the Path topology is compatible with metrical structures varying by a conformal factor, making it more suitable for general, curved geometries. That is, whereas the symmetry group of the Fine topology was the inhomogeneous Lorentz group plus dilations, that of the Path topology is the group of conformal diffeomorphisms on Minkowski spacetime, preserving the null-cone structure.[27]

With so much structure being encoded at the level of topology, and yet not having fixed the metric itself, the stage is set for some kind of best-systems account of metrical structure. As mentioned above, that account could not take the form that it did in special relativity—that is, the symmetry group of the supervenient geometrical structure

[26] The Path topology is also introduced alongside the Fine topology by Guccione [2011, §4.2] and more formally by Naber [2012, Appx. A].

[27] Hawking *et al.* call the former the 'homothecy' group of Minkowski spacetime [p. 176], and call the latter the conformal group. Winnie, however, calls the former the 'Extended Poincaré Group' [1977, p. 182].

could not be pinpointed by the transformations under which the best-systems dynamical laws preserve their form. Instead, perhaps it might be argued that Einstein's equation is part of the best-systematization for a material system including a supervenient metric field, and having a primitive topological structure $\mathcal{P}$. After all, preserving conformal structure would ensure local Lorentz-covariance, leaving only the second of DiSalle's two stipulations to be accounted for by way of a best-systems explanation.

6.3 Summary and Avenues for Further Research

An overview of this thesis's chapters and aims has already been given in the Introduction. But with those chapters now behind us, some additions can be made to that overview, along the lines of what other research projects would naturally follow from the preceding discussions.

After an introduction to the geometrical and dynamical approaches to special relativity (Ch. 1), a critical review was given of recent publications regarding that controversy (Ch. 2). It was argued that Einstein's principle–constructive distinction should not be taken as a comprehensive taxonomy for physical theories and their explanatory power, and that Einstein's distinction has not set the stage well for the ongoing discussion on explanation in relativity. While a brief outline of recent work on scientific explanation was given, there is certainly room for a much larger project along those lines. Scientific explanation is a notoriously complex topic, so that any attempt to more accurately position principle-theories and constructive theories within the spectrum of contemporary models of scientific explanation would be a great undertaking. Although Lorentz's tripartite distinction between 'phenomenological' principle-theories, 'constraint' principle-theories, and 'mechanism' constructive theories is a step in the right direction, it is only one of many in outlining the various forms of theories employed in physics and the kind of explanation or understanding that they provide.

In the remaining chapters, this thesis has focused primarily upon the search for a metaphysical underpinning to the dynamical approach's 'constructivist project', which

stems from the claim that the arrow of explanation runs from dynamics to geometry. As a first step, it was shown that the dynamical approach should be understood as a form of Einstein's 'practical geometry', in which a Humean conception of laws would feature prominently (Ch. 3). It was also shown that something similar seems to have been going on in Moritz Schlick's early work on relativity, especially given his way of establishing a topological structure via his 'method of coincidences'. But this only scratches the surface of Schlick's work on relativity physics and related areas of philosophy. Some, but not all of the many works that Schlick produced during his career (which was so tragically cut short) have been collected [Mulder and van de Velde-Schlick 1979*a*,*b*], while a larger project is currently underway [Rostock 2013]. Developing a more refined understanding of Schlick's views on space and time would be beneficial not only to the history and philosophy of physics in general, but especially to the present project, in particular with regard to the dynamical approach to general relativity.

After tracing similar ideas through Minkowski's and Schlick's works on relativity, some metaphysical underpinnings for the dynamical approach were proposed—namely, those of a relativistic form of regularity relationalism (Ch. 4). It was decided that the standard Euclidean topology could be assumed as the choice of sub-metrical structure, and a case study of the 'constructivist project' was carried out in Chapter 5. But of course the same case-study could be undertaken with alternative choices of sub-metrical structure, or even by considering solutions to different Lorentz-covariant equations. Regarding the choice of sub-metrical structure, it was mentioned in §6.2 that certain topologies, like the 'Path' topology itself, would provide a richer subvening base for the constructivist project without going so far as to encode Minkowski metrical structure itself.[28] And while plenty of work has been done in outlining possible topologies for relativistic spacetimes, the aim of those studies has usually been to encode as much spatiotemporal structure as possible

[28]Thus while it is proposed as a primitive geometrical structure for the dynamical approach to general relativity, the Path topology might also have been included among those considered in §4.2 for the dynamical approach to special relativity. Whereas the Fine topology left only the global scale of the metric to be accounted for by the dynamical laws, the Path topology would leave local scale factors to be so determined, meaning that it would not overshoot the goal quite as badly as the Fine topology did.

at the level of topology, which makes those topologies too rich for the constructivist project.[29] What would be useful to the constructivist project is something that would land somewhere between, say, the Euclidean and Fine topologies. With much less work having been done toward that end, there is surely some low-hanging fruit in the field of developing a 'Goldilocks' topological structure for the dynamical approach to special or general relativity.

Along similar lines, other sub-metrical structures might be considered as potential primitive geometrical structures for the constructivist's Humean mosaic. For example, Appendix C considers the same project as that in Chapter 5, but takes Euclidean affine structure as primitive. And while topological structure was given strong motivation in Chapters 3 and 4, motivations would need to be provided for viewing affine structure as a natural choice of structure for one's Humean mosaic. Or, given that the causal structure of Minkowski spacetime determines its metric only up to a conformal factor, one might consider taking causal relations as primitive instead. Causal structure is easily motivated as a primitive feature of the constructivist's Humean mosaic, although it is less clear that this would leave a sufficient portion of Minkowski geometrical structure to be accounted for by the symmetries of the best-systems dynamical laws.[30] Finally, an alternative primitive structure worth considering comes up in recent work by Tim Maudlin, mentioned above. In a forthcoming, two-volume work [2014*a*; 2014*b*], Maudlin proposes a new foundation for spacetime theories, which he calls the theory of linear structures.[31] In the first volume, Maudlin expresses some concerns with taking topology as a foundation for differential geometry, and argues for linear structure in its stead.[32] Regarding those concerns, there is little more that can be said without going into more detail of Maudlin's

[29] See, for example, the topologies outlined by Guccione [2011]. Other topologies for relativistic spacetime, not mentioned by Guccione, include the s- and t-topologies introduced by Sribatsa Nanda [1972]. (In personal correspondence, Giacomo Dossena has pointed out that Nanda's t-topology is actually coincident with the path topology.) Nanda also offers a procedure for defining topologies with the same symmetry group as Minkowski spacetime [1971].

[30] Recall that the Path topology encodes causal structure. See footnote 28, above.

[31] An outline the ideas in these two volumes is available in a recent publication [2010]. It should be noted that it was to this publication that Dorr was responding on page 161, above.

[32] Some of these concerns are mentioned in Appendix A.

work. But if in fact the theory of linear structures has certain advantages over standard topology, then it should be recalled that the dynamical approach is not necessarily tied to topology as the fundamental spatiotemporal structure. Any naturally motivated sub-metrical structure may do, and perhaps some form of Maudlin-style 'linear structure' should be taken into consideration. At the very least, Maudlin's work can be mined for alternative topologies, to be considered along the lines above. For instance, his proposed linear structure for Minkowski spacetime actually generates a non-standard topology for Minkowski spacetime that is distinct from the standard Euclidean topology $\mathcal{R}^4$ and also the Fine topology $\mathcal{F}$ [Maudlin 2014*b*].

Finally, the current chapter has explored some of the metaphysical aspects of the constructivist project in more detail. The idea of taking topological structure as primitive has been fleshed out in terms of a primitive, first-order, pluralist property, and the possibility of extending the project to richer ontologies of more complicated fields, and perhaps even to the context of general relativity, has been defended. But ending this chapter with so many avenues for further development should not be a distraction from the undertakings of the preceding chapters. The dynamical approach has been defended from recent criticisms, and shown to be in need of a certain metaphysical underpinning in order to be defended from others. Toward that end, it has been shown that the dynamical approach might be construed as a relativistic form of regularity relationalism. And while there may be other ways to interpret and defend the dynamical approach, this one has been shown to be particularly fitting. It has also been shown to be plausible in even some of the simplest physical systems, which of course is the context in which a best-systems account of any structural feature is most likely to struggle.

A Amorphous Geometrical Structures: An Introduction to Topological and Differentiable Structure

In the chapters above, it has been stressed that the arrow of explanation from dynamics to geometry requires that the constructivist assume only some sub-metrical structure, such as topological structure or differentiable structure, as the primitive geometrical structure of the material world. It would therefore be beneficial to review the mathematical representation of these 'amorphous' geometries. Similar accounts can be found in some introductions to the mathematical foundations of spacetime theories, but this section will put an emphasis on the structures most relevant to the chapters above. For consistency with the preceding chapters, this overview is given in terms of a collection P of (field-) points p.

Topological Structure

Topological structure is in some sense the most basic geometrical structure that can be assigned to a set of points.[1] It is the first 'level' of structure through which one must pass in constructing a model of spacetime out of a set P of otherwise unstructured points.[2] Conceptually, topological structure might be summed up in terms of dimensionality and continuity, for it is the structure in virtue of which there come to be matters of fact about whether some collection of the points in P form a closed loop of any given dimension. Along those lines, the properties of an n-dimensional topological space are often illustrated by the behavior of an n-dimensional rubber sheet, which can be stretched and pinched, warped and bent in any way imaginable—just so long as the sheet is not ripped in any way, and so long as no points are 'glued' or 'pasted' together. Tim Maudlin, among others, uses this sort of illustration.

> Suppose you are presented with a collection of qualitatively identical, featureless points and want to organize them into a space. [...] Different sorts

[1] This claim not be confused with the claim that topological structure is more basic, metaphysically speaking, than a higher-level structure like metrical structure. See the discussion in §6.1.

[2] The set P is sometimes referred to as a 'bag' of points.

> of geometrical structure exist, and form a hierarchy. [... And] sitting at the bottom of this definitional hierarchy is a sub-metrical geometry, aspects of a space that do not depend on either the metric or the affinity. This level of geometrical structure is sometimes called 'rubber sheet geometry' because it determines features of a space that remain unchanged if the space is 'stretched' without 'cutting' or 'pasting'. 'Stretching' a space, like stretching a rubber sheet, means changing the distances between points and possibly deforming straight lines into curved lines, but there is still a very rich set of relations that remain invariant. For example, if we were to draw a circle on a rubber sheet and then deform the sheet, the figure would remain a closed curve no matter how the sheet is stretched. This level of geometry is called the *topology* of the space. [Maudlin 2010, pp. 63–4]

But despite its privileged position at the bottom of the hierarchy for geometrical structures, some introductions to the geometrical structures employed in spacetime physics put little emphasis on this rubber-sheet structure. Instead, many jump straight to assigning the points in P with a higher-order geometry, like metrical or pseudo-metrical structure.[3] And when such introductions do mention topology explicitly, it is often in passing. A quotation from the comprehensive text *Gravitation* by Misner, Thorne, and Wheeler illustrates this.[4]

> The mathematician usually begins his development of differential topology by introducing some very primitive concepts, such as sets and topologies of sets, by building a fairly elaborate framework out of them, and by then using that framework to define the concept of a differentiable manifold. But most physicists are satisfied with a more fuzzy, intuitive definition of manifold: roughly speaking, an n-dimensional *differentiable manifold* is a set of "points" tied together continuously and differentiable, so that the point in any sufficiently small region can be put into a one-to-one correspondence with an open set of points of $\mathbb{R}^n$. [Misner *et al.* 1973, p. 241]

But in the present case, since the constructivist takes topological or differentiable structure

[3] E.g. Misner *et al.* [1973, pp. 8–10], Wald [1984, p. 423], or John Norton's introductions to spacetime structure [1992, §II] and the hole argument [2011].

[4] Some other introductions to the mathematical foundations of spacetime theories give a similarly 'fuzzy' introduction to topology [Anderson 1976, Ch. 1; Friedman 1983, §I, §II.1], whereas the word seems not to appear at all in others [Earman 1989]. On the other hand, Sklar [1974, pp. 51–4] gives a more thorough conceptual introduction, and Malament [2012, §1.1] gives a more formal one, following the approach by Geroch [1972, 1989]. As in footnote 6 on page 184, Robert Wald's text [1984] is also an exception.

to be the fundamental, primitive geometrical structure of a material world, it will be beneficial to review the sort of development to which Misner *et al.* allude. And so in what follows, the mathematical formalism of these geometrical structures are outlined. Along the way, we shall see that the formalism is, admittedly, somewhat opaque in regard to concepts like continuity.[5] But at the very least, the outline should clarify the relationship between these amorphous spaces and the 'open sets' used to define them.

Axioms, Open Sets, and Neighborhoods

In contrast to the somewhat conceptual description above, the mathematical definition of a topological space is a set (P) of points with an associated collection ($\mathcal{T}$) of 'open' subsets of that set. What makes the subsets 'open' is simply that they satisfy three axioms.[6]

1. The union of any number of subsets T in $\mathcal{T}$ is also in $\mathcal{T}$.
2. The intersection of any two subsets $T \in \mathcal{T}$ and $U \in \mathcal{T}$ is also in $\mathcal{T}$.
3. The entire set P and the empty set $\emptyset$ are both included in $\mathcal{T}$.

When these axioms are satisfied, the subsets T in $\mathcal{T}$ are deemed to be 'open' subsets of P, which will be expressed here as $T \subseteq P$. The collection $\mathcal{T} = \{T : T \subseteq P\}$ is then called a topology 'on' P, and the topological space itself can be labelled $\langle P, \mathcal{T} \rangle$.[7]

By definition, then, the property of being 'open' is specific to a particular topology: a subset T of P might be open with respect to $\mathcal{T}$, but not with respect to some other topology. And from this we see that the axioms of a topological space do not capture any particular characeteristic of 'openness' that any given subset might have intrinsically. Rather, 'openness' is true of some particular subset primarily in virtue of its relation to other (open) subsets.[8] But this doesn't provide much insight into what 'openness' has

[5] As mentioned in §6.1, this is one of the reasons that Tim Maudlin proposes a new foundation for spacetime structure [2010; 2014*a*].
[6] As listed by Wald [1984, §A]. The outline that follows also draws from Lipschutz [1965].
[7] Incidentally, a 'closed' subset C is defined as the compliment of an open subset. Here, this will be expressed as $C \subseteq P$. Thus both P and $\emptyset$ are not only open, but also closed. A topological space can also be axiomatized in terms of its closed subsets, as discussed for example by Lipschutz [1965, Ch. 5] and Maudlin [2010, pp. 65–6].
[8] This is discussed further in §6.1.

to do with the characteristics of a rubber sheet. For that, a more intuitive description of an open set is sometimes offered, according to which an open set T is one in which from any point $p \in T$, there is some degree of 'wiggle room' available from p before leaving the open set. But of course this description relies upon a picture of open sets embedded in some familiar geometrical space. Indeed, the description owes its origin to the familiar open sets of Euclidean topology, which can be 'induced' by the Euclidian metric as illustrated in Figure 6(a) on page 110, above.[9]

Alternatively, some progress (though arguably not much) in recognizing the continuity structure encoded by open sets comes from using the open sets to define 'neighborhoods' of the points that they contain. A neighborhood of a point p is any subset of P, be it open or closed, that contains some open subset T, which in turn contains p. A neighborhood of some point p is labelled N_p.

$$p \in T \subseteq N_p \subseteq P$$

And from this it follows that open sets themselves are also neighborhoods. Then, for each and every point in the space P, one can define a 'neighborhood system' as the collection of all neighborhoods for p with respect to some particular topology. The neighborhood system of p with respect to the topology $\mathcal{T}$ will be labelled $\mathcal{N}_p^{\mathcal{T}}$.

$$\mathcal{N}_p^{\mathcal{T}} = \{N_p \mid N_p \supseteq T \ni p\}, \quad T \in \mathcal{T}$$

From these definitions there follow certain characteristics of neighborhoods and neighborhood systems in any topological space, and these characteristics can actually be used as alternative axioms of a topological space.[10] Beyond the suggestive nature of their name, neighborhoods do not necessarily conjure up images of a rubber sheet any more quickly than open sets did themselves. However, one way that they do (allegedly) provide some insight into the continuity structure of a topological space is summarized by the following quotation from Aleksandroff.

A topological space is nothing other than a set of arbitrary elements (called

[9] A similar description is discussed by Wald [1984, p. 423] and Lipschutz [1965, p. 66].
[10] This is shown, for instance, by Lipschutz [1965, p. 71]. (*cf.* Baum [1964, p. 20])

> "*points*" of the space) in which a concept of continuity is defined. Now this concept of continuity is based on the existence of relations, which may be defined as local or neighborhood relations—it is precisely these relations which are preserved in a continuous mapping of one figure to another. Therefore, in more precise wording, a topological space is a set in which certain subsets are defined and are associated to the points of the space as their neighborhoods. [...] Using the notion of neighborhood, the concept of continuity can be immediately introduced: A mapping ϕ of a topological space $\langle P, \mathcal{T} \rangle$ onto a (proper or improper) subset of a topological space $\langle X, \mathcal{Y} \rangle$ is called *continuous* at the point p, if for every neighborhood N_x of the point $x = \phi[p]$ one can find a neighborhood N_p of p such that all points of N_p are mapped into points of N_x by means of ϕ. If ϕ is continuous as every point of $\langle P, \mathcal{T} \rangle$, it is called continuous in $\langle P, \mathcal{T} \rangle$. [Alexkandroff 1961, pp. 8–9, with adapted notation]

This at least sheds some light on how the concept of continuity is captured by open-set structure, although it is not entirely clear what work neighborhoods are doing in the process. For while they appear explicitly in Aleksandroff's definition of continuity, it was noted above that open sets themselves are also neighborhoods, and thus many authors refer instead to open sets in giving the above definition of continuity. This is addressed further below.

Topological Equivalence and Homeomorphisms

If the topology $\mathcal{T}$ on P is represented by a collection of open subsets, then a complete description of any given topological structure would call for a list of the subsets in $\mathcal{T}$. Likewise, the differences between various topologies would also boil down to a comparison of their open sets. For extremely simple systems, this would be a feasible project, but more generally it is not. Therefore, in order to compare two topological spaces, and in particular to show whether two topologies are equivalent, alternative methods are required. In short, the two spaces can be shown to be (or not to be) 'homeomorphic', an explanation of which requires the introduction of the following terminology.[11]

[11] What follows mimics the approach by Malament [2012, §1.1]. However, since Malament does not assume a primitive topology on his manifold of spacetime points, but rather 'induces' a topology by assigning the points with differentiable structure, his approach includes a few other possible-mapping axioms that would be redundant here. See also footnote 14 on page 192.

Let $\langle P, \mathcal{T} \rangle$ and $\langle X, \mathcal{Y} \rangle$ be two topological spaces, and consider the possible mappings $\phi : P \to X$ that assign every point $p \in P$ to some point $x \in X$. Call ϕ 'surjective' if it assigns to each point in X at least one point from P. This characteristic is more commonly expressed as ϕ being 'onto'. A similar characteristic is injection: call ϕ 'injective' if no point in X is associated with more than one point in P. Injection is more commonly expressed as ϕ being 'one-to-one'. When ϕ is both injective and surjective (or one-to-one and onto), call it 'bijective'.

If ϕ is bijective, it meets half the requirements of being a homeomorphism. What's left is to establish is bicontinuity. Toward that end, call ϕ 'open' if every open set $T \subseteq P$ is mapped to an open set $Y \subseteq X$. A similar concept is that of continuity, introduced above. Call ϕ 'continuous' if every open set $Y \subseteq X$ is mapped by ϕ^{-1} to an open set $T \subseteq P$. When ϕ is both open and continuous, call it 'bicontinuous'.

Finally, if ϕ is both bijective and bicontinuous, it is a homeomorphism. Such mappings respect, or encode, the topological structure of one space into another other. Thus any two topological spaces are topologically equivalent iff they are homeomorphic—i.e. iff there exists a homeomorphism between them. Any and all topological properties of one of a pair of homeomorphic spaces are shared by the other.

Separation Axioms and Structural Hierarchy

It was mentioned above that if the topology $\mathcal{T}$ on P is represented by a collection of open subsets, then a complete description of any given topological structure would call for a list of the subsets in $\mathcal{T}$, but that for general topological spaces this would be infeasible. For this reason topologies are typically introduced not by listing their open sets, but by stating some further axioms that the open subsets satisfy. These axioms' being satisfied corresponds to a topology's having certain properties, which distinguish it from other topologies. In particular, what 'separation axioms' are satisfied by the open sets of some topological space is what determines how sharply any two points of the space can be distinguished from each other with regard to their topological structure.

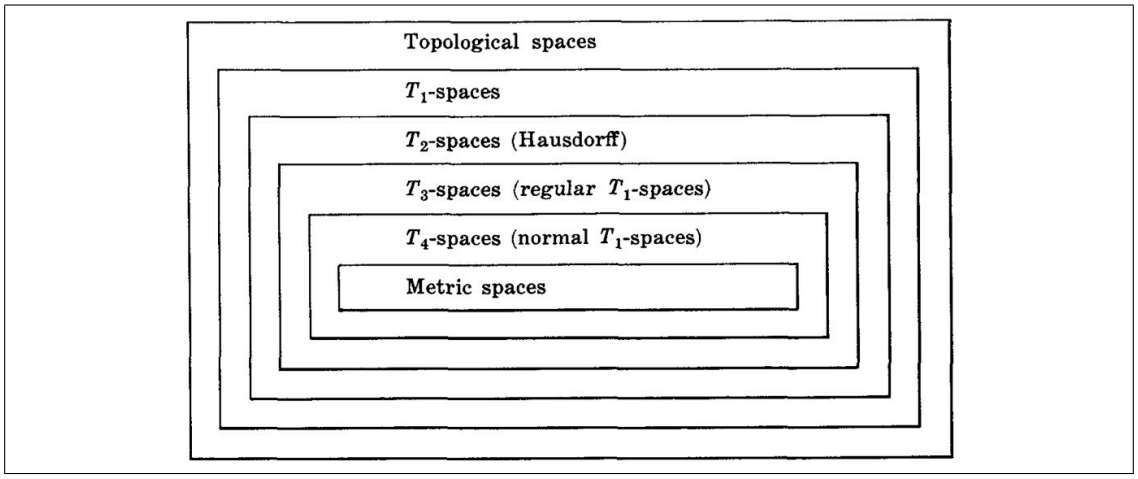


Figure 29: Hierarchy of separation axioms, taken from Lipschutz [1965, p. 141].

There are many separation axioms, but some of the most basic ones are summarized below.[12] The hierarchy of topological spaces satisfying the standard separation axioms is also illustrated in Figure 29.

Before satisfying any separation axioms, the collection of 'open' subsets must satisfy the basic axioms of topology, listed above. Beyond that, a topological space is called 'T_0', or 'Kolmogorov', iff for any two distinct points $p, q \in P$, there exists an open set T that contains only one of them. In this sense, a T_0 space is one in which the topological structure itself is sufficient in discerning between any two points—i.e. it is one in which no two points share the same neighborhood system, and so no two points are topologically indiscernible. More is said on T_0 spaces in Appendix B.

Next on the ladder, a topological space is 'T_1', or 'Fréchet', or 'cofinite' iff for any two numerically distinct points $p, q \in P$, each one belongs to at least one open set that does not also contain the other. And to make this more than trivially possible, a T_1 space is one for which all singleton sets $\{p\}$, and in fact all *finite* subsets, are closed. Thus the points of a T_1 space can be distinguished somewhat more strongly, in terms of the open sets, than they could be in a T_0 space, where only one of the two points must be contained in an open set not containing the other. Necessarily, all T_1 spaces are also T_0 spaces.

[12] This outline draws mainly from Lipschutz [1965, ch 10] and Schechter [1997, ch 16].

pseudometrizable	metrizable
paracompact (partitions of unity)	paracompact and T_0
normal and symmetric (shrinkings and Urysohn functions, plus symmetric)	T_4 = normal and T_1
completely regular (gauges, uniformities)	$T_{3.5}$ = Tychonov (has Hausdorff compactifications)
regular (closed neighborhood bases; extensions by continuity)	T_3 = regular and separated
preregular (limits are unique, up to topological distinguishability)	T_2 = Hausdorff (has unique limits)
symmetric (the closures of points form a partition of X)	T_1 = Fréchet (points are closed)
(arbitrary topological space)	T_0 = Kolmogorov (points are topologically distinguishable)

Figure 30: Given a non-T_0 space that has one of the properties on the left, its Kolmogorov quotient will satisfy the corresponding separation axiom on right. Kolmogorov quotients are introduced in §B, below. This image is taken from Schechter [1997, 1998, §16].

One of the best-known separation axioms comes next in the hierarchy: a topologocal space is 'T_2', or 'Hausdorff', or 'separated', iff for any two numerically distinct points $p, q \in P$, each belongs to an open set that does not contain the other, and those respective open sets are disjoint. Because the open sets mentioned in the description of a T_1 space were not necessarily disjoint, this once again represents a slightly stronger distinction between any two points of the space. And once again, a T_2 space is necesarily also T_1.

Beyond these, the basic hierarchy of separation axioms goes on to include T_3 and T_4 spaces, and even $T_{3.5}$ spaces, which are essentially special cases of T_1 spaces. The details of those categories are less pertinent to the current project, and so needn't be outlined here. They are, however, representative of the structural properties that lie 'between' a Hausdorff space and a metric space, as illustrated in Figure 29. Finally, because Appendix B gives special attention to T_0 spaces, it should be noted that various other axioms can still be met by the open sets of a space that fails to satisfy even those requirements. The table in Figure 30 summarizes the properties of non-Kolmogorov spaces and their Kolmogorov counterparts.[13]

[13]Non-T_0 spaces are oftentimes not addressed in introductions to point-set topology. Eric Schechter's book is an exception [1997, pp. 435–8].

Differentiable Structure

If topological structure sits at the bottom of the geometrical hierarchy, then differentiable structure is the next up. For whereas topological structure encodes continuity, and perhaps a certain concept of 'nearness', differentiable structure puts a restriction on the continuity structure by requiring that it be 'smooth' to some specified degree. This is a particularly important feature of any geometrical structure to be employed in spacetime physics, as the use of differential equations assumes a certain smoothness structure to the space being coordinatized. Cian Dorr summarizes this well.

> The standard mathematical apparatus used for stating physical theories about space-time is that of differential geometry. Mostly, everything is done on the assumption that space-time forms a *smooth manifold*. Smooth manifolds are mathematical objects somewhat richer than mere topological spaces: we can think of them as capturing the structure of a slightly less amorphous kind of rubber sheet, which can be deformed only by gentle stretches and squeezes which never introduce anything like a 'kink' or 'corner'. [Dorr 2011, p. 148]

This 'smoothness' structure can be encoded by a particular collection, or 'atlas', of coordinatizations of a topological space—mappings from portions of the space into richer structures like $\mathbb{R}^4$. But before outlining that process, it is worth noting that differentiable structure cannot be ascribed to just any topological space: all talk of differentiable structure is restricted to what are called 'topological manifolds'. A topological space $\langle P, \mathcal{T} \rangle$ is a topological manifold $M_{\mathcal{T}}$ iff it is both T_2 and locally homeomorphic to a Euclidean space. The latter restriction amounts to two requirements: first, that for each point $p \in P$, there is a neighborhood N_p that is homeomorphic to some open subset $R \subseteq \mathcal{R}^4$, where $\mathcal{R}^4$ is the standard 'induced' topology of $\mathbb{R}^4$ (Fig. 6(a), p. 110); second, that any pair of these local homeomorphisms $\phi_{_N} : N_p \to R$ are 'pair-wise compatible' with each other, meaning that either their respective domains do not intersect, or else their compositions are 'autohomeomorphisms' from $\mathbb{R}^4$ to $\mathbb{R}^4$. These requirements are illustrated in Figure 31.

The distinction between a topological space and a topological manifold is critical only

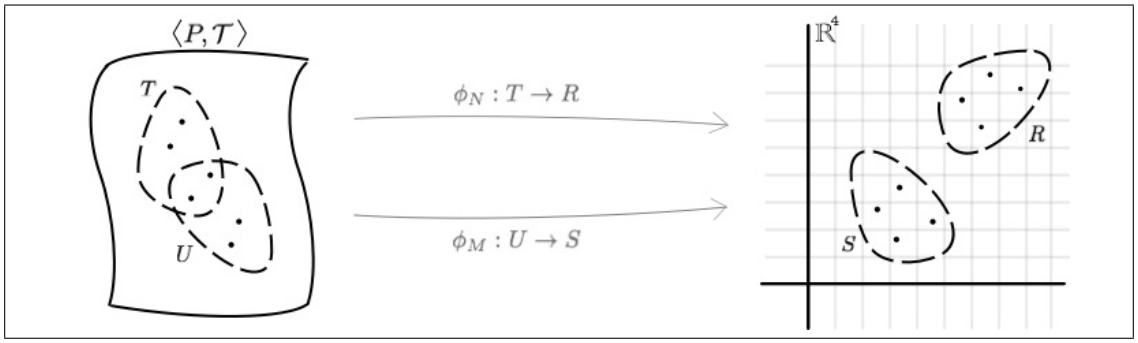


Figure 31: Two local homeomorphisms ϕ_N and ϕ_M that map open subsets of $\langle P, \mathcal{T} \rangle$ into open subsets of $\mathbb{R}^4$ with the standard Euclidean topology. If $\phi_N \circ \phi_M^{-1}$ and $\phi_M \circ \phi_N^{-1}$ are autohomeomorphisms of $\mathbb{R}^4$ onto itself, then the two mappings are 'compatible'. (Alternatively, if $T \cap U = \emptyset$, then the two mappings are automatically compatible.) A set $\mathcal{D}$ of compatible local mappings that cover $\langle P, \mathcal{T} \rangle$ is called an 'atlas'.

because of the fact mentioned above—that differentiable structure is defined in terms of a topological space's possible mappings into $\mathbb{R}^4$. Consider, for instance, Malament's description of a differentiable manifold.

> A smooth n-dimensional manifold ($n \geq 1$) can be thought of as a point set to which has been added the "local smoothness structure" of $\mathbb{R}^n$.
>
> [Malament 2012, p. 1]

And this "local smoothness structure" can be assigned to a topological manifold in terms of the autohomeomorphisms mentioned above. Quite simply, it can be required that the ϕ_N are 'pair-wise compatible' to some degree (C^k), meaning that if their domains intersect, then the partial derivatives, up to some order k, are all continuous. And in this case, the topological manifold $M_{\mathcal{T}}$ is called a 'differentiable manifold' $M_{\mathcal{D}}$. Its topology $\mathcal{T}$ is unchanged, and the collection of pair-wise compatible, k-continuous local homeomorphisms ϕ_N is called an 'atlas' $\mathcal{D}$. It is in virtue of the fact that the ϕ_N in $\mathcal{D}$ are compatible to some degree k that their compositions (also called 'transition maps') are called 'diffeomorphisms' and that the space itself is called 'differentiable'. Furthermore, the degree of their compatibility determines the degree of the space's smoothness: the transition maps of a C^k atlas have well-defined k^{th} partial derivatives. An atlas must be at least C^1 for the space $\langle P, \mathcal{T}, \mathcal{D} \rangle$ to be a differentiable manifold $M_{\mathcal{D}}$, and the term 'smooth' is reserved for those differentiable manifolds for which the transition maps are continuous to all orders—i.e. for those with a C^∞ atlas.

In this way, differentiable structure can be put in its place in the hierarchy alongside the topological structures introduced above. To say that $\langle P, \mathcal{T} \rangle$ is a topological space is to say that the sets in $\mathcal{T}$ satisfy the axioms of topology. To say that $\langle P, \mathcal{T} \rangle$ is a topological manifold is to say that it is a Hausdorff (T_2) space that is locally homeomorphic to $\mathbb{R}^4$, in that it admits a collection (atlas) $\mathcal{D}$ of mappings (charts) $\phi_{_N} : N_p \to R$ that cover $\langle P, \mathcal{T} \rangle$ and have compositions (transition maps) $\phi_{_N}[\phi_{_M}^{-1}] : \mathbb{R}^4 \to \mathbb{R}^4$ that are all autohomeomorphisms (C^0). Then, to say that $\langle P, \mathcal{T}, \mathcal{D} \rangle$ is a differentiable or smooth manifold is to say that the transition maps mentioned above are actually diffeomorphisms to some degree (C^k or C^∞, respectively). And finally, because a differentiable manifold is by definition a topological space, the notation $\langle P, \mathcal{T}, \mathcal{D} \rangle$ is partially redundant. A differentiable or smooth manifold might simply be written $\langle P, \mathcal{D} \rangle$, or even just $M_{\mathcal{D}}$.[14]

[14] As suggested in footnote 11 on p. 186, Malament's recent work illustrates that redundancy [2012, pp. 4–5].

B Multiple Fields and Material Events: Reducing Coincidence to Topological Indiscernibility

In Chapter 5, solutions to the Klein–Gordon equation were explored as potentially supportive examples of a set of scalar field values which, when assigned the Euclidean topological structure 'induced' by Minkowski spacetime, are best systematized by a Lorentz-covariant equation. However, if more than one physical field is to be taken as primitive in more realistic systems, this raises the question as to whether the constructivist's ontology and ideology would need to be expanded. A discussion regarding potential changes to the ontology can be found in Chapter 6. In this appendix, the focus is upon changes to the ideology. In particular, if 'material events' are to be understood as collections of coincident field values, it would seem that by including multiple physical fields in the fundamental ontology, the constructivist must also include in the fundamental ideology some primitive coincidence relation between the values of distinct physical fields.[1] But in what follows, it will be shown that for certain classes of topological structures, this coincidence relation between field values can be made redundant. In short, coincidence can be reduced to topological indiscernibility.[2]

Material Events as Coincident 'Field-Points'

Let the dynamical approach to special relativity, understood as a form of regularity relationalism, assume the existence of a finite number (m) of distinct physical fields, each comprising point-like parts. Call the point-like parts 'field-points', and consider the entire collection of the world's m fundamental fields' points, ignoring for the moment any primitive spatiotemporal structure. As in the chapters above, let this fundamental ontology be represented by a set P of field-points p, q, r, etc. Of what physical field any given field-point is a part can be represented by its belonging to one of m mutually disjoint subsets $F_i \subset P$, each of which comprise the point-like parts of one particular

[1] This definition of material events is introduced in the discussion surrounding footnote 25 on page 88. It is discussed further below.

[2] In this section, familiarity with the notation and terminology introduced in Appendix A is assumed.

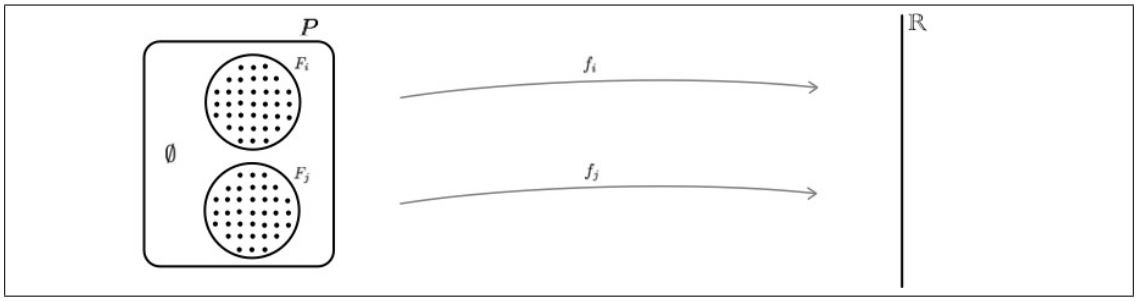


Figure 32: A set P of field-points having degrees of freedom represented by functions f_i and f_j into $\mathbb{R}$, and belonging to distinct fields F_i and F_j, the sum of which is P. Thus in this example, $m = 2$. Note that in this figure, no geometrical structure of the field-points is illustrated.

physical field, so that $\bigcup_i F_i = P$. The degrees of freedom for the field-points of each particular field can be represented by mappings $f_i : F_i \to I$, with I being an interval in $\mathbb{R}$.[3] This kind of representation is illustrated in Figure 32.

When values of numerous physical fields are co-instantiated, we have the concept of a 'material event'. Consider Brown's aforementioned notion of material events: "the set of values at a point of the (components of the) most fundamental fields in our best physical theories" [2005*b*, p. 13]. By this definition, material events can be reduced to field-points if a "set of values" is interpreted as a collection of field-points standing in some coincidence relation, which will be expressed here as $\mathrm{R_c}$. For example, in a world with two fundamental physical fields F_i and F_j, if $p \in F_i$ and $q \in F_j$ and $p\mathrm{R_c}q$, then p and q constitute an event $\textsc{e} = \{p, q\}$.

Now, for material events to be so construed, the coincidence relation $\mathrm{R_c}$ must have particular features. First, it should be clear that the coincidence relation $\mathrm{R_c}$ must have the properties of an equivalence relation.

$$\text{Reflexivity: } \forall p(p \in P \to p\mathrm{R_c}p)$$
$$\text{Symmetry: } \forall p \forall q(p\mathrm{R_c}q \to q\mathrm{R_c}p)$$
$$\text{Transitivity: } \forall p \forall q \forall r(p\mathrm{R_c}q \wedge q\mathrm{R_c}r \to p\mathrm{R_c}r)$$

Second, it must be the case that no material event E own more than one field-point

[3]For simplicity, only scalar fields are considered here. See also the discussion in §6.2. And as mentioned in §6.1, it is assumed that fields do indeed exist where their values are 0.

belonging to any particular F_i.[4]

$$\forall p \forall q \forall i (p, q \in F_i \ \wedge \ p \mathrm{R_c} q \ \rightarrow \ p = q)$$

From this, and the fact that field-points are taken to exist even when $f_i[p] = 0$, it follows that every material event comprises exactly m coincident field-points.

Before continuing, it should be noted that in the absence of any further structure, any realistic example of F_i will include some numerically distinct yet qualitatively indiscernible field-points. That is to say, each F_i will own more than one pair of points (p, q) such that while p and q share all the same (non-haecceitist) properties, $p \neq q$.[5] And as a result, some material events will comprise field-points that are numerically distinct yet qualitatively indiscernible from the field-points that constitute some other material event, and such material events will themselves be indiscernible. For example, in the same hypothetical world of two physical fields F_i and F_j, consider the following four field-points: $p, q \in F_i$ and $r, s \in F_j$. Suppose that $f_i[p] = f_i[q]$ and $f_j[r] = f_j[s]$, and suppose also that $p \mathrm{R_c} r$ and $q \mathrm{R_c} s$. Then in the absence of any structural or haecceitist properties, the distinct events $\textsc{e} = \{p, r\}$ and $\textsc{f} = \{q, s\}$ will be intrinsically indiscernible. This is the issue that Brown was addressing when he assumed material events to have some kind of primitive arrangement, as described at the close of §3.2, above. But here again, the question of whether topological structure can in fact individuate such material events can be put aside for the time being.[6] Instead, what is of interest now is whether topological structure can be ascribed to the fundamental field-points in a way that would make redundant the coincidence relation held by any set of m field points that constitute a material event.

Coincidence as Topological Indiscernibility

Assuming a coincidence relation $\mathrm{R_c}$ on the world's collection P of field-points defines a collection E of material events $\textsc{e}$, each of which is an equivalence class of the coinci-

[4] Throughout this section, $=$ is used to refer to numerical identity, not mere qualitative indiscernibility.
[5] That is to say, points for which the principle of the identity of indiscernibles fails when the quantifier is restricted to non-haecceitist properties.
[6] This is discussed in footnote 11 on page 104.

dence relation $\mathrm{R_c}$. This means that when multiple physical fields are included in the constructivist's ontology, the constructivist project could actually be discussed in terms of material events, all the while understanding structural claims about material events to be translatable into structural claims about the fundamental field-points. But in order to be allowed this privilege, the constructivist needn't take the coincidence relation as part of the fundamental ideology. Rather, the coincidence relation $\mathrm{R_c}$ can be construed as topological indiscernibility. For if field-points have no more geometrical structure than mere topology, then to call two field-points 'coincident' carries no more weight than to say that the two points are topologically indiscernible. Thus the constructivist could purge coincidence relations from the fundamental ideology by assuming instead that the collection P of field-points has a primitive topology that renders certain collections of m field-points (of mutually distinct fields) topologically indiscernible. In turn, material events themselves would be reduced to collections of topologically indiscernible field-points.

To clarify: to take as primitive some topological structure for a collection P of field-points is to take as primitive a topological space $\langle P, \mathcal{T} \rangle$, where the topology on P is represented by a set $\mathcal{T}$ of 'open subsets' T.[7]

$$\mathcal{T} = \{T \mid T \subseteq P\}$$

Those open subsets can be used to assign a neighborhood system $\mathcal{N}_p$ to any given field-point p.

$$\mathcal{N}_p = \{N_p \mid p \in T \subseteq N_p \subseteq P\}$$

Therefore, to say that the field-points' topological structure $\mathcal{T}$ renders certain collections of field-points 'topologically indiscernible' is simply to say that not all field-points have

[7]The basics of point-set topology are outlined in Appendix A. As explained there, the symbols $\subseteq$ and $\subseteq$ represent open and closed subset relations, respectively.

unique neighborhood systems.[8]

$$\exists p \exists q (p \neq q \ \wedge \ \mathcal{N}_p = \mathcal{N}_q)$$

And in topological jargon, this all amounts to the assumption that $\langle P, \mathcal{T} \rangle$ is not 'T_0'—i.e. that $\langle P, \mathcal{T} \rangle$ is not a 'Kolmogorov' space. For if it were, then for every pair of points p, q, there would be at least one open set T containing either p or q but not the other.[9] And so it is by assuming $\mathcal{T}$ to be a non-Kolmogorov topology that allows coincidence to be reduced to the having of identical neighborhood systems. And since having identical neighborhood systems is an equivalence relation, this reduction is in accord with the requirements on R_c outlined above.

Working with $\langle E, \mathcal{Q} \rangle$ Instead of $\langle P, \mathcal{T} \rangle$

It was suggested above that if material events are construed as collections of topologically indiscernible field-points, then the constructivist project could described strictly in terms of material events. But to be sure, it must be shown that if $\langle P, \mathcal{T} \rangle$ is a non-Kolmogorov topological space, then the set E of material events E also forms a topological space, and that there is some formal relationship between the topology $\mathcal{T}$ on P and the topology on the set E of material events E. Only then could it safely be said that any claims about the topological structure of material events can be translated into claims about the topological structure of the field-points themselves. Fortunately, it can be shown that whatever non-Kolmogorov topology is assigned to the field-points themselves, there is a one-to-one correspondence between that topology and the topology on the space of material events. Put another way, the topological structure of the material events can be determined uniquely by a non-Kolmogorov topology $\mathcal{T}$ on P. This details are outlined below, and the idea is illustrated in Figure 33.

[8] And of course if collections of topologically indiscernible field-points are to represent material events, then the field-points having the same neighborhood systems must be of mutually distinct fields.
[9] The hierarchy of separation axioms for topological spaces is included in Appendix A.

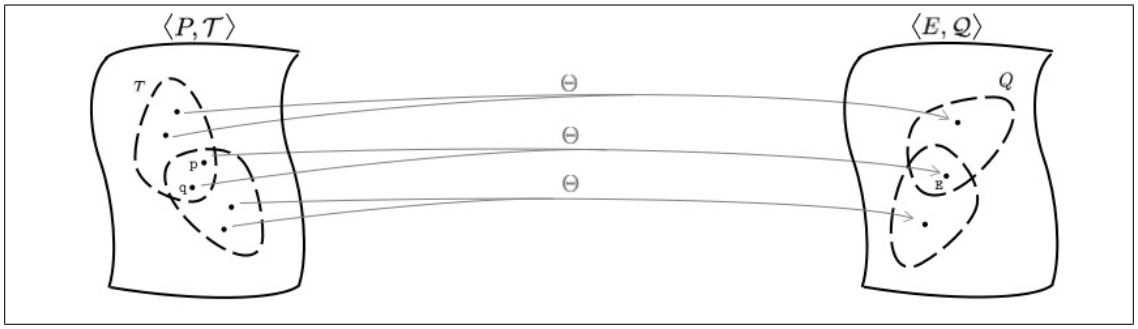


Figure 33: A hypothetically non-Kolmogorov topological space $\langle P, \mathcal{T} \rangle$, and a Kolmogorov-quotient mapping Θ from $\langle P, \mathcal{T} \rangle$ into the quotient topological space $\langle E, \mathcal{Q} \rangle$. Θ maps equivalence classes of the coincidence relation $\mathrm{R_c}$ to single points in $\langle E, \mathcal{Q} \rangle$, which is by definition a Kolmogorov space. For instance, p and q are illustrated as having the same neighborhood system, and are therefore mapped to E.

First, recall that material events are equivalence classes of topological indiscernibility.

$$\mathrm{E} = [p]_{\mathcal{N}} = \{p\} \cup \{q \mid \mathcal{N}_q = \mathcal{N}_p\}$$

And because the having of identical neighborhood systems is an equivalence relation, we could also express this in terms of a coincidence relation.

$$\mathrm{E} = [p]_{\mathrm{R}_c} = \{p\} \cup \{q \mid p \mathrm{R_c} q\}$$

Next, consider the collection E of such material events.

$$E = \{\mathrm{E} \mid \mathrm{E} = [p]_{\mathrm{R}_c},\ p \in P\}$$

This set is referred to as the 'quotient' of P with respect to the equivalence relation $\mathrm{R_c}$.

$$E = P/\mathrm{R_c}$$

Then, the topology $\mathcal{T}$ on P can be used to identify a unique 'quotient' topology $\mathcal{Q}$ on E, the open sets of which are simply the quotients of the open sets T with respect to the equivalence relation $\mathrm{R_c}$.

$$\mathcal{Q} = \{Q \mid Q = T/\mathrm{R_c},\ T \in \mathcal{T}\}$$

Thus the topological space $\langle E, \mathcal{Q} \rangle$ can itself be called the 'quotient' of $\langle P, \mathcal{T} \rangle$ with respect

to the equivalence relation.

$$\langle E, \mathcal{Q} \rangle = \langle P, \mathcal{T} \rangle / \mathrm{R_c}$$

By definition, none of the points in the quotient space will have the same neighborhood system, and the quotient space is therefore a Kolmogorov (T_0) space. Thus $\langle E, \mathcal{Q} \rangle$ is called the 'Kolmogorov quotient' of $\langle P, \mathcal{T} \rangle$. The quotient space can be thought of as the space that would result from 'gluing' together the members of each set of topologically indiscernible points in $\langle P, \mathcal{T} \rangle$.[10]

What has been established is a 'quotient map' $\Theta : P \to E$, which maps points in P to their equivalence classes in E. While Θ is not injective, it is surjective and also bicontinuous: Θ maps open sets of $\mathcal{T}$ into open sets of $\mathcal{Q}$, and Θ^{-1} maps the latter into the former. And while Θ is not bijective itself, its application in this context does establish a bijection between the open sets of $\langle P, \mathcal{T} \rangle$ and those of $\langle E, \mathcal{Q} \rangle$.[11] Thus any claims made about $\langle E, \mathcal{Q} \rangle$ can be translated into claims about $\langle P, \mathcal{T} \rangle$, meaning that the constructivist project could indeed by carried out in terms of material events, all the while avoiding any commitment to a primitive coincidence relation.

[10] Relationships between non-T_0 spaces and their Kolmogorov quotients are outlined in Figure 30 (p. 189).
[11] [Schechter 1997, p. 437]. Note that this does not entail a bijection between the set of possible non-Kolmogorov topologies on P and the set of possible Kolmogorov topologies on P.

C Another Constructivist Project: Scalar Fields with Primitive Affine Structure

In §4.2, it was pointed out that if the dynamical approach assigns only a Euclidean topological structure to the point-like parts of physical fields, then the constructivist project involves a somewhat bolder best-systems claim than that of Huggett's Leibnizian relationalist. Seeing that the constructivist's best-systems claim would be on surer footing if some richer geometrical structure were taken as primitive, a less symmetric sub-metrical structure is considered in this appendix—namely, the affine structure of Minkowski spacetime, which is the standard affine structure of $\mathbb{R}^4$. This structure lands between topological and metrical structure in the standard hierarchy of spacetime geometries, meaning that while it amounts to more geometrical structure than $\mathcal{R}^4$, it still leaves metrical structure to supervene upon the symmetries of the dynamical laws in a best-systems account of Minkowski geometry.[1]

So, whereas in Chapter 5 it was the standard topological structure of $\mathbb{R}^4$ being taken as primitive, in this section it will be the standard affine structure of $\mathbb{R}^4$ that is taken as primitive.[2] And while the former was labelled $\mathcal{T}$, the latter will be labelled A.[3] Otherwise, the approach of this section will mirror that outlined in §4.3, and undertaken in §5.2. First, the question at hand will be briefly re-introduced in terms of this affine structure A. Then, it will be argued that by taking A as primitive, the best-systems account of Minkowski geometry is on very sure footing even in the case of a single, relatively simple scalar field.

The New Question at Hand

In this context, the question at hand is whether there any examples of a four-dimensional, scalar-field solution to a Lorentz-covariant equation, such that the solution could not, under an appropriate *affine transformation*, solve some otherwise-covariant equation with

[1] Friedman, among others, describes the place of affine structure in the standard hierarchy [1983, pp. 10–2].
[2] Thanks are due to Teruji Thomas for suggesting this approach.
[3] This should not be confused with the Alexandrov topology $\mathcal{A}$, which was discussed in §4.2.

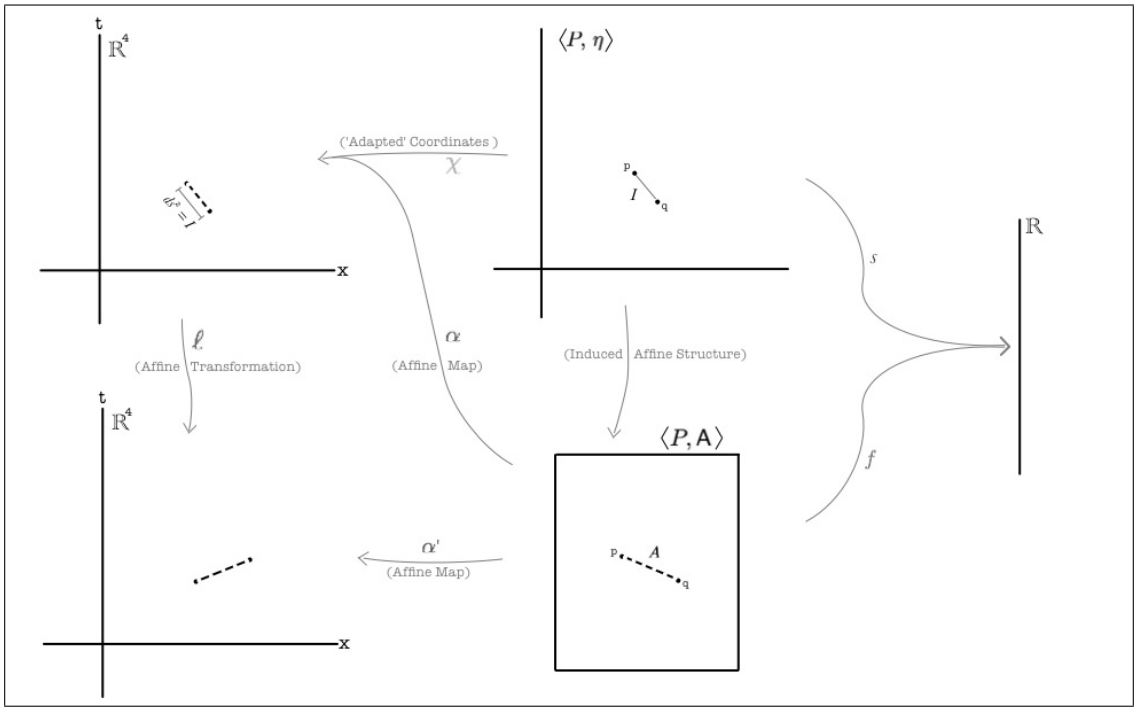


Figure 34: When a physical field $\langle P, \eta \rangle$ is coordinatized so that $ds^2 = I$, the field points' affine structure coincides with that of $\mathbb{R}^4$. The space $\langle P, \mathsf{A} \rangle$ is thus a potentially supportive example for this version of the constructivist project. It can be coordinatized by the affine mapping $\alpha : \langle P, \mathsf{A} \rangle \to \mathbb{R}^4$ as well as by any affine mapping $\alpha' = \ell \circ \alpha$, where $\ell : \mathbb{R}^4 \to \mathbb{R}^4$ is an affine transformation.

a greater or equal balance of simplicity and strength. Such a scalar field, when considered 'up to affinity', would serve as a concrete example for this version of the constructivist project. To spell the question out in more detail, the discussion in §4.3 should be rehearsed and tailored to Euclidean affine structure. This discussion is accompanied by Figure 34, which is a modification of Figure 9 on page 116.

As before, consider some solution ψ to a particular Lorentz-covariant equation E. Also as before, think of the solution ψ as describing a physical scalar field, understood as a set of field points having scalar-valued degrees of freedom and standing in Minkowski interval relations.[4] Just as in Figure 9, Figure 34 models the field points' degrees of freedom by a mapping $s : \langle P, \eta \rangle \to \mathbb{R}$, and also models their metrical structure by the 'adapted' coordinatization $\chi : \langle P, \eta \rangle \to \mathbb{R}^4$. Thus the solution $\psi(t, \bar{x})$ can once again be modeled by a composition of these two mappings, so that $\psi = s \circ \chi^{-1}$.

[4]As in footnote 24 (p. 115), this interpretation takes physical fields to be entities in their own right.

Here begin the differences between the question associated with Figure 9 and that with Figure 34. In the former case, it was a four-dimensional, topologically structured entity being taken as primitive, and the topology structure $\mathcal{T}$ was chosen to be the one 'induced' upon the field points by the choice of a Euclidean metric on an adapted coordinatization χ. Thus the starting point was to assume the existence of the set P of field points that constitute a solution to some Lorentz-covariant equation E, and then to endow to those field points with that induced structure $\mathcal{T}$. But now a four-dimensional, *affinely* structured entity is being taken as primitive. Thus the starting point will be to assume the existence of the set P of field points that constitute a solution to some Lorentz-covariant equation E, and then to endow those field points with the induced Euclidean affine structure A.[5] The resulting affine space $\langle P, \mathsf{A}\rangle$ is the referent of the phrase, "a solution ψ, considered up to affinity". Using these terms, the question at hand is whether there are any examples of a Lorentz-covariant equation E with solutions such that $\langle P, \mathsf{A}\rangle$ cannot be systematized by any other equations with a different covariance group than E, and with a greater or equal balance of simplicity and strength than E. A few more words about Figure 34 will clarify what is meant by systematization in this context.

Just as in Figure 9, the values of the field points in $\langle P, \mathsf{A}\rangle$ can be modeled by a mapping $f : \langle P, \mathsf{A}\rangle \to \mathbb{R}$, and their Euclidean affine structure A can be modeled by an affine mapping into $\mathbb{R}^4$. Let the affine mapping $\alpha : \langle P, \mathsf{A}\rangle \to \mathbb{R}^4$ map the field points to the same positions in $\mathbb{R}^4$ as in the aforementioned adapted coordinatization χ. Thus the values of the field points can be described by the coordinate-dependent function $f_\alpha(t, \bar{x})$, where $f_\alpha = f \circ \alpha^{-1}$. There is of course a wide range of coordinate-dependent descriptions of the field points' values: one for every possible coordinatization of $\langle P, \mathsf{A}\rangle$. But once again, when it comes to systematizing the space $\langle P, \mathsf{A}\rangle$, only a subset of those possible coordinatizations will be permissible—in this case, those which 'respect', or 'preserve'

[5] In the case of an induced topological structure $\mathcal{T}$, the open sets T were generated by the 'open balls' of the Euclidean metric (§4.2). In the case of an induced affine structure A, the straight lines A are those deemed by the Euclidean metric to be the shortest paths between pairs of points.

the field points' Euclidean affine structure A by being affine mappings between $\langle P, \mathsf{A}\rangle$ and $\mathbb{R}^4$. In Figure 34, the subset of permissible alternative coordinatizations for $\langle P, \mathsf{A}\rangle$ can be described as the collection that includes any coordinatization α' that results from an affine transformation $\ell : \mathbb{R}^4 \to \mathbb{R}^4$, so that $\alpha' = \ell \circ \alpha$. Therefore the set of permissible, coordinate-dependent descriptions of the field points' values will comprise f_α and any $f_{\alpha'}$ associated with some such α'.

Simplifying the Question at Hand

Because the goal—a best-systems account of Minkowski geometry—remains the same, it will again be useful to conduct a case-study on the Klein–Gordon equation. And for simplicity, the following discussion will be restricted to its real-valued solutions.[6] Now the question at hand becomes whether there are any examples of a real-valued solution $\psi_{\Re}$ to the Klein–Gordon equation which, when considered 'up to affinity' as $\langle P, \mathsf{A}\rangle$, admits no permissible alternative coordinatizations α' with respect to which $f_{\alpha'}$ solves an equation with a greater balance of simplicity and strength than the Klein–Gordon equation itself. And one advantage of having taken affine structure as primitive is that it becomes trivial to map out the various forms of the $f_{\alpha'}$ associated with all permissible coordinatizations of $\langle P, \mathsf{A}\rangle$. Indeed, it is a handy feature of affine transformations that, when applied to a (real-valued) planewave solution ψ (or $\psi_{\Re}$), the result is another planewave.

To see this, consider the four-dimensional description of a planewave, where $\vec{r} = (t, \bar{x})$ and $\vec{\varkappa} = (-\omega, \bar{\kappa})$ and '$\cdot$' represents a four-dimensional Euclidean dot-product.[7]

$$\psi(\vec{r}) = Ae^{i(\vec{\varkappa}\cdot\vec{r})}$$

The result of applying an affine coordinate transformation $\ell = M\vec{r} + \vec{a}$, with M being a linear transformation matrix and $\vec{a}$ being a four-vector, would be the following.

$$\psi(\vec{r}) \xrightarrow{\ell} Ae^{i(\vec{\varkappa}\cdot(M\vec{r}+\vec{a}))}$$

[6] For this reason, the 'rival' systematizations under consideration will be those introduced in §5.2.
[7] In Chapter 5, planewaves were described in Equation 5 (p. 124) as functions of t and $\bar{x}$: $\psi = Ae^{i(\bar{\kappa}\cdot\bar{x}-\omega t)}$.

The terms can then be re-arranged, using $M^\intercal$ as the transpose of M.

$$\psi(\vec{r}) \xrightarrow{\ell} Ae^{i(\vec{\varkappa}\cdot\vec{a})}e^{i(M^\intercal\vec{\varkappa}\cdot\vec{r})} \tag{17}$$

Sure enough, Equation 17 is a planewave with (four-dimensional) wave-vector $M^\intercal\vec{\varkappa}$ and a phase shift of $e^{i(\vec{\varkappa}\cdot\vec{a})}$. It follows that for any $\psi_{\Re}$ of the form of Equation 8 (p. 125), the permissible affine transformations ℓ will produce recoordinatizations α' according to which $f_{\alpha'}$ also takes that same form. For comparison, this can also be expressed in the original, three-vector notation, using ϑ to represent a constant phase shift.

$$f_\alpha = \Re\left(Ae^{i(\bar{\kappa}\cdot\bar{x}-\omega t)}\right) \xrightarrow{\ell} f_{\alpha'} = \Re\left(Ae^{i(\bar{\kappa}'\cdot\bar{x}-\omega' t+\vartheta)}\right)$$

And so it is that in considering a real-valued planewave solution to the Klein–Gordon equation 'up to affinity' as $\langle P, \mathsf{A}\rangle$, it is trivial to list all forms that the $f_{\alpha'}$ can take: the only option is Equation 8 (p. 125). This greatly simplifies the question at hand; all else that must be determined is what other equations might be solved by such $f_{\alpha'}$. If any, those are the rival systematizations of $\langle P, \mathsf{A}\rangle$.

Illustrating Affine Transformations on Planewaves

To proceed, it will be useful to recall what makes a planewave a planewave, and then to distinguish those characteristics from what makes a planewave solve some particular equation. (The lesson above is that only the latter is affected by affine transformations.) As mentioned in §5.1, and as illustrated in Figure 11 (p. 125), a real-valued planewave is a wave with constant amplitude A, constant angular frequency ω, constant wavenumber κ (and thus phase speed v_p), and constant magnitude in all directions orthogonal to the wave-vector $\bar{\kappa}$. These constants vary from planewave to planewave, and even vary among the planewave solutions to a particular equation. However, in the latter case, certain restrictions hold on the constant values of angular frequency and wavenumber. Indeed, *what it is* for a planewave to solve a particular equation is for its angular frequency

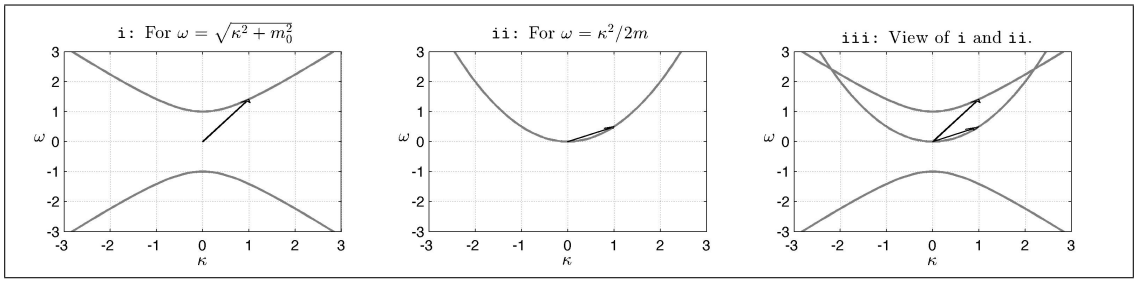


Figure 35: Plots i and ii show the dispersion relation (in grey) for solutions to the Klein–Gordon equation and Schrödinger (or Squaringer) equation, respectively. Plot iii shows both. For each, a two-dimensional wave-vector $\vec{\varkappa}$ is shown (in black) for the choice of $m = \kappa = 1$.

and wavenumber to satisfy these restrictions corresponding to that particular equation.[8] For instance, it was explained in §5.1 that a planewave solution to the Klein–Gordon equation must satisfy its dispersion relation: $\omega = \sqrt{\kappa^2 + m_0^2}$, whereas a planewave solution to the Schrödinger (or Squaringer) equation must satisfy another: $\omega = \kappa^2/2m$. Thus the salient distinctive features of a pair of planewave solutions to two different equations are captured in the relationships between ω and κ. And when these values are written as the components of a *four*-dimensional wave-vector $\vec{\varkappa} = (-\omega, \bar{\kappa})$, these distinctive features can be illustrated as in Figure 35.[9] There, what makes it the case that some planewave solves a particular equation is for that planewave's wave-vector $\vec{\varkappa}$ to have components represented by a point on the corresponding curve.

Now, the usefulness of these diagrams becomes clear in the connection between affine transformations ℓ on $\langle P, \mathsf{A} \rangle$ and coordinate transformations on the diagrams themselves. That connection is as follows: an affine transformation ℓ on the x, t coordinates of a planewave solution to some equation E will transform it into a planewave solution of another equation F if and only if there exists a linear transformation which, if applied to the κ, ω coordinates of the planewave's wave-vector diagram, would transform the wave-vector into one that satisfies the dispersion relation of equation F. To see this, note that a stretching of the κ-coordinate in a wave-vector diagram effects a decrease in κ, which amounts to an increase in wavelength, which is effected by a *contraction* of the x-coordinate in a planewave plot. Meanwhile, a stretching of the ω coordinate in a

[8] See footnote 5 on page 124, footnote 8 on page 127, and the discussion surrounding them.
[9] Thanks are due to Teruji Thomas for suggesting this illustration.

wave-vector diagram corresponds to a stretching of the t coordinate in the corresponding planewave's plot. Or alternatively, simply recall from above (Eqn. 17, p. 206) that an affine transformation maps planewaves to planewaves by first applying the transpose of the linear transformation matrix M on the four-dimensional wave-vector $\vec{\varkappa}$, and then adding a phase shift. Because phase shifts do not affect the dispersion relation of a planewave, they can be ignored. Thus in order to determine whether some given $\langle P, \mathsf{A} \rangle$ can be coordinatized in such a way that $f_{\alpha'}$ solves some other equation, it need only be determined whether there is a linear transformation on its four-dimensional wave-vector diagram that would transform the wave-vector into one whose components satisfy some other equations' dispersion relation.

Finally, it should be noted that all of the above has been said with regard to a single planewave solution to the Klein–Gordon equation. But of course any linear combination of planewave solutions to the Klein–Gordon equation is also a solution. Fortunately, the simplification outlined above applies in this case as well. If affine transformations take planewaves to planewaves, then they also take linear combinations of planewaves to linear combinations of planewaves. So, if a linear combination of n real-valued planewave solutions to the Klein–Gordon equation is being taken up to affinity as $\langle P, \mathsf{A} \rangle$, the question becomes whether there is a single linear transformation that would transform each of the summands' wave-vectors into ones that satisfy the dispersion relation of some one other equation.[10]

Systematizing Solutions to the Klein–Gordon Equation

With the stage set, let the games begin. To start, one might consider the simple case of a single, real-valued planewave solution to the Klein–Gordon equation, as illustrated in Figure 12-i (p. 128). For some such $\langle P, \mathsf{A} \rangle$, it is not difficult to see that there is indeed a permissible coordinatization α' according to which $f_{\alpha'}$ solves another equation with

[10] This begs the following question: if some rival equation F is solved by some $f_{\alpha'}$, then is F also solved by each of the (transformed) planewave summands? That the answer is yes follows from the linear independence of the planewave summands.

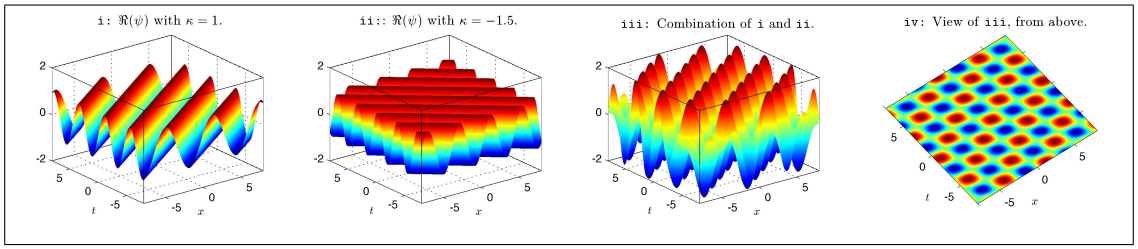


Figure 36: The sum of two real-valued planewave solutions to the Klein–Gordon equation, for which $\omega = \sqrt{\kappa^2 + m_0^2}$. As above, $A = m_0 = \hbar = 1$. The difference in the summands' wavenumbers κ ensures that their sum is not a planewave. The two summand planewaves' wave-vectors are shown in Figure 37.

a different covariance group. A quick look at Figure 35-iii shows that there are linear transformations M under which such a planewave's wave-vector would come to satisfy the dispersion relation of, say, the Squaringer equation. For example, consider some simple stretching of the ω-axis. Indeed, the discussion surrounding Figure 12 (p. 128) already made it clear that under a corresponding stretching of the t-axis, a single, real-valued planewave solution to the Klein–Gordon equation could be transformed into a solution of the Squaringer equation. Therefore it seems that the $\langle P, \mathsf{A} \rangle$ under consideration is just as easily systematized by at least one rival equation with a different covariance group.

Moving beyond the very simplest examples, one might then consider a linear combination of two real-valued planewave solutions to the Klein–Gordon equation, having different wavenumbers κ. The resulting $\langle P, \mathsf{A} \rangle$ is only slightly less symmetric than either of the planewave summands themselves, as was illustrated in Figure 17 (p. 138). A similar example, which will serve better in the present context, is illustrated in Figure 36. But here again, while the transformation is not quite as obvious, there is indeed a single linear transformation M under which both of the summands' wave-vectors would come to satisfy the dispersion relation of, say, the Squaringer equation. As shown in Figure 37, a simple stretching of the ω coordinate would transform one of the two wave-vectors appropriately, and the second wave-vector could then be appropriately transformed—without affecting the first—via a shearing transformation. Indeed, this linear transformation on the wave-vectors corresponds to the same coordinate transformation (on the planewaves themselves) that was illustrated in Figure 19 (p. 140). And so here again, it could not be argued that the $\langle P, \mathsf{A} \rangle$ under consideration is arguably best-systematized by a Lorentz-covariant

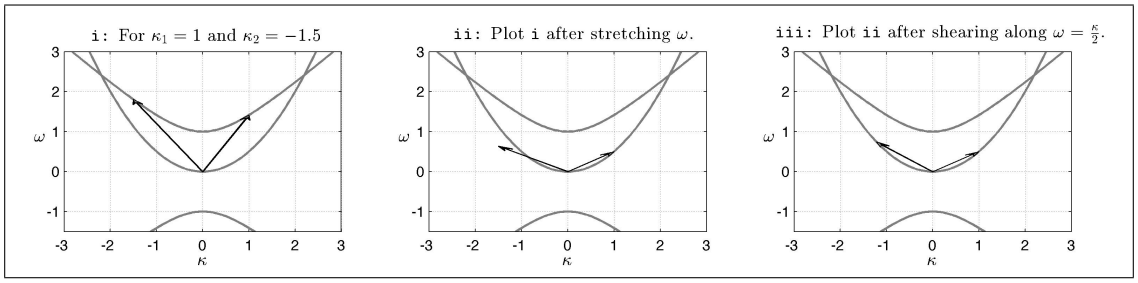


Figure 37: Plot i shows the wave-vectors for two planewave solutions to the Klein–Gordon equation, which are illustrated in Figure 36. Plot ii shows the same wave-vectors after stretching the ω-axis, so that the rightmost wave-vector comes to satisfy the dispersion relation of the Squaringer equation. Plot iii shows the result after then shearing the coordinates along the line of the rightmost wave-vector.

equation.

However, when considering the $\langle P, \mathsf{A} \rangle$ that results from the sum or three or more planewave solutions to the Klein–Gordon equation, all having unique wavenumbers, the rival systematizations begin to be eliminated. For example, in order for the sum of three or more such planewaves to be recoordinatized so that $f_{\alpha'}$ now solves, say, the Squaringer equation, there must be some linear transformation M under which three, non-collinear points on a hyperbola are transformed into three non-collinear points on a parabola. But of course linear transformations take hyperbolae to hyperbolae, and so under any affine transformation, the three summands' wave-vectors will still have endpoints landing on a hyperbola. As will be emphasized below, this means that under any permissible coordinatization α', the corresponding mathematical description $f_{\alpha'}$ of some such $\langle P, \mathsf{A} \rangle$ will still solve an equation with a hyperbolic dispersion relation. In order to also solve another equation with some different dispersion relation, such as the parabolic one of the Squaringer equation, the three transformed wave-vectors would have to have endpoints that fall on that corresponding curve as well.

In fact, there are circumstances under which this is possible. For example, two-dimensional conic sections like hyperbolae and parabolae can intersect up to four times. This means that if the transformed wave-vectors' endpoints coincided with the points of intersection, then the corresponding $f_{\alpha'}$ would indeed solve the Squaringer equation. But of course if these conic sections intersect in at most four places, it follows that for the $\langle P, \mathsf{A} \rangle$

resulting from the sum of five or more planewave solutions to the Klein–Gordon equation, all having different wavenumbers, there will be no single linear transformation M under which each of the summands' wave-vectors would come to satisfy the dispersion relation of the Schrödinger equation. To be clear, this means that under no affine transformation ℓ would the corresponding $f_{\alpha'}$ solve the Squaringer equation. Thus for even a relatively simple solution to the Klein–Gordon equation, such as a linear combination of 5 or more planewaves, Euclidean affine structure is rich enough to impede systematization by equations with parabolic dispersion relations.

To take the argument further, the same must be shown for equations of any other dispersion relation. Fortunately, the result is not hard to generalize. Being restricted to affine transformations, the planewave summands for any solution to the Klein–Gordon equation will always have wave-vectors that fall on a hyperbola, meaning that under any permissible coordinatization α', the corresponding $f_{\alpha'}$ will solve only those equations with a hyperbolic dispersion relation, or at least with one that intersects a hyperbola at least as many times as the number of summands. By considering examples of solutions that result from a linear combination of higher numbers of planewave summands, the latter possibility can be ruled out on the grounds of simplicity. As for the former, the onus would be on the critic of the constructivist project to find a non-Lorentz-covariant equation with a hyperbolic dispersion relation, the simplicity and strength of which rival that of the Klein–Gordon equation. The difficulty in this task becomes clear when one considers that a hyperbolic dispersion relation between ω and κ comes about naturally from even-ordered, partially differential equations in which the temporal and spatial derivative terms are of equal order and opposite sign, and in which a separate constant (mass) term also appears—like the Klein–Gordon equation.[11] What other forms of equations could have a hyperbolic dispersion relation is not clear.

[11] Compare, for instance, the dispersion relations of the very similar 'rival' equations outlined in §5.2.

Bibliography (Author Index)

Alexkandroff, Paul (1961). *Elementary concepts of topology.* Dover Publications.
Cited on page(s) 186.

Anderson, James (1976). *Principles of relativity physics.* Academic Press.
Cited on page(s) 23, 183.

Arntzenius, Frank (2012). *Space, time, and stuff [electronic resource].* Oxford University Press: Oxford Scholarship Online.
Cited on page(s) 95, 166, 169.

Arons, A. B. and Peppard, M. B. (1965, May). Einstein's proposal of the photon concept—a translation of the Annalen der Physik paper of 1905. *American Journal of Physics*, **33**(5), 367–374.
Cited on page(s) 216.

Balashov, Yuri and Janssen, Michel (2003, Jun). Review: Presentism and relativity. *British Journal for the Philosophy of Science*, **54**(2), 327–346. Reviewed works: *The tensed theory of time: A critical examination*, by William Lane Craig.
Cited on page(s) 21, 23, 29, 48, 49, 51, 52.

Barbour, Julian (1982). Relational concepts of space and time. *The British Journal for the Philosophy of Science*, **33**(3), 251–274.
Cited on page(s) 89.

Baum, John (1991 [1964]). *Elements of point set topology.* Dover Publications, Inc.
Cited on page(s) 185.

Beebee, Helen (2000, Nov). The nongoverning conception of laws of nature. *Philosophy and Phenomenological Research*, **61**(3), 571–594.
Cited on page(s) 77.

Bell, John (1987). How to teach special relativity. In *Speakable and unspeakable in quantum mechanics: Collected papers in quantum philosophy*, Chapter 9, pp. 67–80. Cambridge University Press. Reprinted from *Progress in Scientific Culture*, 1:2, 1976.
Cited on page(s) 25, 27.

Bell, John (1992, Sep). George Francis Fitzgerald. *Physics World*, **5**(9), 31–35. 1989 lecture; abridged by Denis Weare.
Cited on page(s) 26.

Belot, Gordon (2011). *Geometric possibility.* Oxford University Press.
Cited on page(s) 74, 95.

Bradley, James (1727). A letter from the reverend Mr. James Bradley, Savilian professor of astronomy at Oxford, and F.R.S. to Dr. Edmond Halley astronom. reg. &c. giving an account of a new discovered motion of the fix'd stars. *Philosophical Transactions of the Royal Society of London*, **36**, 637–661.
Cited on page(s) 7.

Bromberger, Sylvian (1966). Why-questions. In *Mind and cosmos* (ed. R. Colodny), Volume 3 of *University of Pittsburgh Series in the Philosophy of Science Series*. University of Pittsburgh Press.
Cited on page(s) 41.

Brown, Harvey (1993). Correspondence, invariance and heuristics in the emergence of special relativity. In *Correspondence, invariance and heuristics: Essays in honour of Heinz Post* (ed. S. French and H. Kamminga), Volume 148 of *Boston Studies in the Philosophy of Science Series*, pp. 227–260. Kluwer Academic Publishers.
Cited on page(s) 30.

Brown, Harvey (1997). On the role of special relativity in general relativity. *International Studies in the Philosophy of Science*, **11**(1), 67–81.
Cited on page(s) 30, 88, 172.

Brown, Harvey (2003). Michelson, FitzGerald and Lorentz: The origins of relativity revisited. *Bulletin de la Société des Sciences et des Lettres de Łódź*, **53**(39), 23–55. A preprint is also available on the PhilSci Archive: #987.
Cited on page(s) 25, 30.

Brown, Harvey (2005*a*, Nov). Einstein's misgivings about his 1905 formulation of special relativity. *European Journal of Physics*, **26**(6), S85–90.
Cited on page(s) 10.

Brown, Harvey (2005*b*). *Physical relativity: Space-time structure from a dynamical perspective*. Oxford University Press.
Cited on page(s) 7, 9, 10, 12, 13, 14, 15, 19, 20, 21, 25, 30, 32, 33, 38, 53, 54, 55, 59, 63, 65, 69, 80, 87, 88, 89, 90, 99, 158, 165, 174, 195.

Brown, Harvey (2009, Nov). The behaviour of rods and clocks in general relativity, and the meaning of the metric field. e-print: arXiv:0911.4440v1 [gr-qc].
Cited on page(s) 59.

Brown, Harvey and Holland, Peter (1999, Mar). The Galilean covariance of quantum mechanics in the case of external fields. *American Journal of Physics*, **67**(3), 204–214.
Cited on page(s) 123.

Brown, Harvey and Lehmkuhl, Dennis (TBD). Einstein, the reality of space, and the action–reaction principle. In *The nature of reality* (ed. P. Ghose). Forthcoming by Oxford University Press. Preprint available on the PhilSci Archive: #9792.
Cited on page(s) 83.

Brown, Harvey and Pooley, Oliver (2001). The origin of the spacetime metric: Bell's 'Lorentzian pedagogy' and its significance in general relativity. In *Physics meets philosophy at the Planck scale: Contemporary theories in quantum gravity* (ed. C. Callender and N. Huggett), Chapter 11, pp. 256–272. Cambridge University Press.
Cited on page(s) 9, 14, 25, 30, 31, 32, 33, 172, 173.

Brown, Harvey and Pooley, Oliver (2006). Minkowski space-time: A glorious non-entity. In *The Ontology of Spacetime* (ed. D. Dieks), Volume 1 of *Philosophy and Foundations of Physics Series*, pp. 67–89. International Society for the Advanced Study of Spacetime: Elsevier.
Cited on page(s) 14, 21, 25, 29, 30, 31, 32, 33, 48, 51, 52, 53, 54, 55, 65, 78, 80, 90, 165.

Brown, Harvey and Sypel, Roland (1995). On the meaning of the relativity principle and other symmetries. *International Studies in the Philosophy of Science*, **9**(3), 235–253.
Cited on page(s) 30.

Brown, Harvey and Timpson, Christopher (2006). Why special relativity should not be a template for a fundamental reformulation of quantum mechanics. In *Physical theory and its interpretation [electronic resource]: Essays in honor of Jeffrey Bub* (ed. W. Demopoulos and I. Pitowsky), Volume 72 of *The Western Ontario Series in Philosophy of Science Series*, Chapter 2, pp. 29–42. Springer.
Cited on page(s) 14, 30, 38, 54.

Buchwald, Diana (ed.) (2002). *The collected papers of Albert Einstein*, Volume 7: The Berlin Years: Writings, 1918–1921. Princeton University Press.
Cited on page(s) 217.

Burgess, John (2014). Lewis on set theory. Forthcoming in *The Blackwell Companion to David Lewis* (ed. B. Loewer and J. Schaffer). Preprint available online: <http://www.princeton.edu/~jburgess/Lewis.pdf>.
Cited on page(s) 162.

Butterfield, Jeremy, Hogarth, Mark, and Belot, Gordon (ed.) (1996). *Spacetime*, Volume 17 of *The International Research Library of Philosophy: Metaphysics and Epistemology Series*. Dartmouth. Series Editor: John Skorupski.
Cited on page(s) 227.

Cameron, P., Hocking, J. G., and Naimpally, S. A. (1974, August–September). Nearness—A better approach to continuity and limits. *American Mathematical Monthly*, **81**(7), 739–745.
Cited on page(s) 103.

Carroll, John (2011). Laws of nature. In *The Stanford Encyclopedia of Philosophy* (Summer 2011 edn) (ed. E. Zalta). The Metaphysics Research Lab: Center for the Study of Language and Information, Stanford University.
Cited on page(s) 77.

Catoni, Francesco, Boccaletti, Dino, Cannata, Roberto, Catoni, Vincenzo, Nichelatti, Enrico, and Zampetti, Paola (2008). *The mathematics of Minkowski space-time [electronic resource]: With an introduction to commutative hypercomplex numbers*. Frontiers in Mathematics Series. Birkhäuser.
Cited on page(s) 17.

Catoni, Francesco, Boccaletti, Dino, Cannata, Roberto, Catoni, Vincenzo, and Zampetti, Paola (2011). *Geometry of Minkowski space-time [electronic resource]*. SpringerBriefs in Physics Series. Springer.
Cited on page(s) 17.

Craig, William Lane (2000*a*). *The tensed theory of time: a critical examination*, Volume 293 of *Synthese Library Series*. Kluwer Academic.
Cited on page(s) 27.

Craig, William Lane (2000*b*). *The tenseless theory of time: a critical examination*, Volume 294 of *Synthese Library Series*. Kluwer Academic.
Cited on page(s) 27.

Craig, William Lane (2001). *Time and the metaphysics of relativity*. Number 84 in Philosophical Studies Series. Kluwer Academic.
Cited on page(s) 5, 9, 10, 23, 27, 28, 29, 48, 52, 79.

Craig, William Lane (2008). The metaphysics of special relativity: three views. In *Einstein, relativity, and absolute simultaneity* (ed. W. L. Craig and Q. Smith), Chapter 1, pp. 11–49. Routledge.
Cited on page(s) 5, 28, 29, 79.

de Senarmont, Henri, Verdet, Émile, and Fresnel, Léonor (ed.) (1866–1870). *Oeuvres complètes d'Augustin Fresnel*, Volume 1–3. Imprimerie Impériale.
Cited on page(s) 6.

Dieks, Dennis (2009). Understanding in physics: Bottom-up versus top-down. In *Scientific Understanding: Philosophical Perspectives* (ed. H. de Regt, S. Leonelli, and K. Eigner), pp. 230–248. University of Pittsburgh Press.
Cited on page(s) 34, 44, 81.

Dieks, Dennis (2010). The adolescence of relativity: Einstein, Minkowski, and the philosophy of space and time. See Petkov [2010], Chapter 9, pp. 225–245.
Cited on page(s) 82.

DiSalle, Robert (1995, May). Spacetime theory as physical geometry. *Erkenntnis*, **42**(3), 317–337.
Cited on page(s) 91.

DiSalle, Robert (2006). *Understanding space-time: The philosophical development of physics from Newton to Einstein*. Cambridge University Press.
Cited on page(s) 34, 39, 176.

Dorato, Mauro (2007, Mar). Relativity theory between structural and dynamical explanations. *International Studies in the Philosophy of Science*, **21**(1), 95–102.
Cited on page(s) 30.

Dorato, Mauro and Felline, Laura (2010). Structural explanations in Minkowski spacetime: Which account of models? In *Space, Time, and Spacetime* (ed. V. Petkov), Volume 167, Part 2 of *Fundamental Theories of Physics Series*, pp. 193–207. Springer.
Cited on page(s) 21, 50.

Dorr, Cian (2010, Sep). Of numbers and electrons. *Proceedings of the Aristotelian Society*, **110**(2), 133–181.
Cited on page(s) 95, 164, 165.

Dorr, Cian (2011, Apr). The geometry of space-time II: Physical geometry and fundamental metaphysics. *Proceedings of the Aristotelian Society*, **111**(1, Part 1), 135–159.
Cited on page(s) 66, 95, 161, 166, 190.

Dorr, Cian and Hawthorne, John (2013). Naturalness. In *Oxford studies in metaphysics* (ed. K. Bennett and D. Zimmerman), Volume 8, pp. 3–75. Oxford University Press.
Cited on page(s) 104.

Dougherty, John (2013). Topology in Lewis's "New Work". Unpublished manuscript, sent in personal correspondence.
Cited on page(s) 103.

Earman, John (1989). *World enough and space-time: Absolute versus relational theories of space and time*. MIT Press.
Cited on page(s) 16, 20, 24, 65, 94, 157, 183.

Eddington, Arthur (1923). *The mathematical theory of relativity* (1st edn). Cambridge University Press.
Cited on page(s) 174.

Einstein, Albert (1905*a*). Über einen die erzeugung und verwandlung des lichtes betreffenden heruistischen gesichtspunkt. *Annalen der Physik*, **17**(1), 132–148. Note: all quotations and page numbers are taken from the English translation [Arons and Peppard 1965].
Cited on page(s) 14.

Einstein, Albert (1905*b*). Zur elektrodynamik bewegter körper. *Annalen der Physik*, **322**(10), 891–921. Note: all quotations and page numbers are taken from the English translation [Einstein 1952*b*]. Other English translations are provided by Stachel [1998, pp. 123–160] and Miller [1981, pp. 392–415].
Cited on page(s) 5, 10, 11, 12, 19.

Einstein, Albert (1907*a*). Bermungen zu der notiz von Hrn. Paul Ehrenfest: "Die translation deformierbarer Elektronen und der Flächensatz". *Annalen der Physik*, **328**(6), 206–208. Note: all quotations and page numbers are taken from the English translation [Stachel 1989*a*, Doc. 44, pp. 236–237].
Cited on page(s) 15.

Einstein, Albert (1907*b*). Ueber das Relativitätsprinzip und die aus demselben gezogenen Folgerungen. *Jahrbuch der Radioaktivität und Elektronik*, **4**, 411–462.
Cited on page(s) 13.

Einstein, Albert (1916, Mar). Die grundlage der allgemeine relativitätstheorie. *Annalen der Physik*, **49**(7), 769–822. Note: all quotations and page numbers are taken from the English translation [Einstein 1952*a*]. A more recent English translation is also available [Kox *et al.* 1997, Doc. 30, pp. 147–200].
Cited on page(s) 85, 218.

Einstein, Albert (1919). My theory. *The London Times*, **28 November**, 13–14. Note: all quotations and page numbers are taken from the version in *Ideas and Opinions* [Einstein 1954*b*], which has also been reprinted elsewhere [Buchwald 2002, Doc. 25, pp. 100–105].
Cited on page(s) 14, 36, 37, 218.

Einstein, Albert (1921). *Geometrie und erfahrung: Erweiterte fassung des festvortrages gehalten an der preussischen akademie der wissenschaften zu berlin, am 27 Januar 1921.* Berlin: Julius Springer. Note: all quotations and page numbers are taken from the English translation in *Ideas and Opinions* [Einstein 1954*a*].
Cited on page(s) 70, 71, 72, 73, 74, 83, 84, 218.

Einstein, Albert (1924). Review of Alfred C. Elsbach, Kant und Einstein. *Deutsche Literaturzeitung*, **1**(24), 1685–1692. Note: all quotations and page numbers are taken from the English translation, forthcoming in Volume 14 of the Collected Papers of Albert Einstein.
Cited on page(s) 73, 74.

Einstein, Albert (1925). Nichteuklidische geometrie und physik [Non-euclidean geometry and physics]. *Die Neue Rundschau*, **36**(1), 16–20. Translation forthcoming in Vol. 14 of the Collected Papers of Albert Einstein.
Cited on page(s) 70, 73, 74.

Einstein, Albert (1936, Mar). Physics and reality. *Journal of the Franklin Institute*, **221**(3), 348–382. Note: all quotations and page numbers are taken from the reprinting in *Ideas and Opinions* [Seelig 1954, pp. 290–323].
Cited on page(s) 37.

Einstein, Albert (1948). Time, space, and gravitation. In *Out of my later years*, Chapter 11, pp. 54–58. Thames and Hudson.
Cited on page(s) 37.

Einstein, Albert (1949*a*). Autobiographical notes. In *Albert Einstein: Philosopher-Scientist* (1st edn) (ed. P. Schilpp), Volume VII of *The Library of Living Philosophers Series*. Cambridge University Press.
Cited on page(s) 11, 14, 15, 58.

Einstein, Albert (1949*b*). Remarks to the essays appearing in this collective volume. See Schilpp [1949], Chapter III, pp. 663–687.
Cited on page(s) 73, 74, 75.

Einstein, Albert (1952*a*). The foundation of the general theory of relativity. See Perrett and Jeffery [1952], Chapter VII, pp. 109–164. English translation of the original publication [Einstein 1916].
Cited on page(s) 217.

Einstein, Albert (1952*b*). On the electrodynamics of moving bodies. See Perrett and Jeffery [1952], Chapter 3, pp. 35–65. An unabridged and unaltered republication of the translation first published by Methuen and Company, Ltd. in 1923.
Cited on page(s) 216.

Einstein, Albert (1954*a*). Geometry and experience. See Seelig [1954], pp. 232–246. Originally a lecture before the Prussian Academy of Sciences, January 27, 1921. First published in German [Einstein 1921].
Cited on page(s) 217.

Einstein, Albert (1954*b*). What is the theory of relativity? See Seelig [1954], pp. 227–232. Originally published in the *London Times* [Einstein 1919].
Cited on page(s) 217.

Einstein, Albert (1954 [1914]). Principles of theoretical physics. See Seelig [1954], pp. 220–223. Inaugural address before the Prussian Academy of Sciences.
Cited on page(s) 37.

Elster, Charlotte (2013, Oct). URL: <http://www.phy.ohiou.edu/~elster/>.
Cited on page(s) 143.

Felline, Laura (2011, Dec). Scientific explanation between principle and constructive theories. *Philosophy of Science*, **78**(5), 989–1000.
Cited on page(s) 34.

Field, Hartry (1980). *Science without numbers: A defense of nominalism*. Oxford: Blackwell.
Cited on page(s) 157.

Field, Hartry (1984). Can we dispense with space-time? In *Proceedings of the 1984 biennial meeting of the Philosophy of Science Association* (ed. P. Asquith and P. Kitcher), Volume 2, pp. 33–90. Philosophy of Science Association: University of Chicago Press.
Cited on page(s) 157, 158.

FitzGerald, George (1889, May). The ether and the Earth's atmosphere. *Science*, **13**(328), 390.
Cited on page(s) 7, 8.

Flores, Francisco (1999). Einstein's theory of theories and types of theoretical explanation. *International Studies in the Philosophy of Science*, **13**(2), 123–134.
Cited on page(s) 38, 39, 44.

Friedman, Michael (1974, Jan). Explanation and scientific understanding. *Journal of Philosophy*, **71**(1), 5–19.
Cited on page(s) 42.

Friedman, Michael (1983). *Foundations of space-time theories: Relativistic physics and philosophy of science*. Princeton University Press.
Cited on page(s) 9, 10, 17, 19, 20, 22, 23, 57, 86, 183, 202.

Friedman, Michael (2002). Geometry as a branch of physics: Background and context for Einstein's 'geometry and experience'. In *Reading natural philosophy: Essays in the history and philosophy of science* (ed. D. Malament), pp. 193–229. Chicago: Open Court.
Cited on page(s) 20, 72, 73, 75, 83, 84, 85, 86.

Frisch, Mathias (2005, Dec). Mechanisms, principles, and Lorentz's cautious realism. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **36**(4), 659–679.
Cited on page(s) 46.

Frisch, Mathias (2011, Aug). Principle or constructive relativity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **42**(3), 176–183.
Cited on page(s) 34, 45, 46, 50, 53, 54, 55, 59, 65, 78.

Galison, Peter (2003). *Einstein's clocks, Poincaré's maps: Empires of time*. Sceptre.
Cited on page(s) 38.

Geroch, Robert (1972). Differential geometry. Unpublished notes, available online: <http://home.uchicago.edu/geroch/Links_to_Notes.html>.
Cited on page(s) 183.

Geroch, Robert (1978). *General relativity from A to B*. University of Chicago Press.
Cited on page(s) 87.

Geroch, Robert (1989, Oct). Manifolds. Unpublished notes, available online: <http://home.uchicago.edu/geroch/Links_to_Notes.html>.
Cited on page(s) 183.

Giovanelli, Marco (2012, Dec). Talking at cross-purposes: How Einstein and logical positivists never agreed on what they were disagreeing about. *Synthese*, **Published Online**, 1–45.
Cited on page(s) 72, 74, 83.

Giulini, Domenico (2006). Algebraic and geometric structures in special relativity. In *Special relativity: Will it survive the next 101 years?* (ed. J. Ehlers and C. Lämmerzahl), Number 702 in Lecture Notes in Physics Series, Chapter 4, pp. 45–111. Springer.
Cited on page(s) 17.

Giulini, Domenico (2010). The rich structure of Minkowski spacetime. See Petkov [2010], Chapter 4, pp. 83–132.
Cited on page(s) 17, 111.

Glazebrook, Richard (1928, Feb). Prof H. A. Lorentz, for. mem. R.S. *Nature*, **121**(3043), 287–290.
Cited on page(s) 10.

Göbel, Rüdiger (1976*a*, May). The smoothpath topology for curved space–time which incorporates the conformal structure and analytic Feynman tracks. *Journal of Mathematical Physics*, **17**(5), 845–853.
Cited on page(s) 176.

Göbel, Rüdiger (1976*b*). Zeeman topologies on space-times of general relativity theory. *Communications in Mathematical Physics*, **46**(3), 289–307.
Cited on page(s) 176.

Göbel, Rüdiger (1980). Natural topologies on Lorentzian manifolds. *Mitteilungen der Mathematischen Gesellschaft*, **10**(8), 763–771. Occasionally cited as: *Topologies for curved space-time defined by paths*, Berichte des Math. Seminars Hamburg 10 (1980), pp. 763–771.
Cited on page(s) 176.

Goldstein, Herbert, Poole, Charles, and Safko, John (2002). *Classical mechanics* (3rd edn). Addison–Wesley Publishing Company, Inc.
Cited on page(s) 122.

Guccione, Giovanni (2011, Sep). Shaping a spacetime from causal structure. Master's thesis, Imperial College, London. Supervisor: Helen Fay Dowker.
Cited on page(s) 177, 180.

Hagar, Amit (2008, Sep). Length matters: The Einstein–Swann correspondence and the constructive approach to the special theory of relativity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **39**(3), 532–556.
Cited on page(s) 54.

Hall, Ned (2012). David Lewis's metaphysics. In *The Stanford Encyclopedia of Philosophy* (Fall 2012 edn) (ed. E. Zalta). The Metaphysics Research Lab: Center for the Study of Language and Information, Stanford University.
Cited on page(s) 104.

Hawking, Stephen and Ellis, George (2010 [1973]). *The large scale structure of space-time [electronic resource]*. Cambridge Monographs on Mathematical Physics Series. Cambridge University Press.
Cited on page(s) 111.

Hawking, Stephen, King, A., and McCarthy, P. (1976, Feb). A new topology for curved space-time which incorporates the causal, differential, and conformal structures. *Journal of Mathematical Physics*, **17**(2), 174–182.
Cited on page(s) 176, 177.

Heaviside, Oliver (1888). The electro-magnetic effects of a moving charge. *Electrician*, **22**, 147–148.
Cited on page(s) 7, 27.

Hemmo, Meir and Hagar, Amit (2013, Aug). The primacy of geometry. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **44**(3), 357–364.
Cited on page(s) 62, 63.

Hempel, Carl (1945, Jan). Geometry and empirical science. *The American Mathematical Monthly*, **52**(1), 7–17.
Cited on page(s) 70, 75.

Hempel, Carl (1965). *Aspects of scientific explanation, and other essays in the philosophy of science*. Free Press; Collier-Macmillan.
Cited on page(s) 40, 221.

Hempel, Carl and Oppenheim, Paul (1948, Apr). Studies in the logic of explanation. *Philosophy of Science*, **15**(2), 135–175. Reprinted elsewhere by Hempel [1965].
Cited on page(s) 40, 41.

Hilbert, David (1971). *Foundations of Geometry* (2nd English edn). Open Court: La Salle, Illinois. Translated by Leo Unger from the 10th german edition of *Grundlagen der Geometrie*. Revised and enlarged by Paul Bernays.
Cited on page(s) 83.

Hoefer, Carl (1996, Jan). The metaphysics of space-time substantivalism. *Journal of Philosophy*, **93**(1), 5–27.
Cited on page(s) 158.

Howard, Don (1984). Realism and conventionalism in Einstein's philosophy of science: The Einstein–Schlick correspondence. *Philosophia Naturalis*, **21**, 616–629.
Cited on page(s) 83.

Howard, Don (2010). Einstein's philosophy of science. In *The Stanford Encyclopedia of Philosophy* (Summer 2010 edn) (ed. E. Zalta). Stanford University.
Cited on page(s) 15.

Howard, Don (2014). Einstein and the development of twentieth-century philosophy of science. Forthcoming in *The Cambridge Companion to Einstein* (ed. M. Janssen and C. Lehner).
Cited on page(s) 72, 74, 83.

Huggett, Nick (1999). Why manifold substantivalism is probably not a consequence of classical mechanics. *International Studies in the Philosophy of Science*, **13**(1), 17–34.
Cited on page(s) 95.

Huggett, Nick (2006, Jan). The regularity account of relational spacetime. *Mind*, **115**(457), 41–73.
Cited on page(s) 94, 97, 98, 165, 167.

Huggett, Nick (2009, Jul). Essay review: *Physical Relativity* and *Understanding Space-Time*. *Philosophy of Science*, **36**(3), 404–422.
Cited on page(s) 30, 64, 65, 90, 106, 113.

Hughes, R. I. G. (1989). *The structure and interpretation of quantum mechanics*. Harvard University Press.
Cited on page(s) 21.

Hughes, R. I. G. (1993, Sep). Theoretical explanation. *Midwest Studies In Philosophy*, **18**(1), 132–153.
Cited on page(s) 21.

Janssen, Michel (1995). *A comparison between Lorentz's ether theory and special relativity in the light of the experiments of Trouton and Noble*. Ph. D. thesis, University of Pittsburgh.
Cited on page(s) 50.

Janssen, Michel (2002*a*). COI stories: Explanation and evidence in the history of science. *Perspectives on Science*, **10**(4), 457–522.
Cited on page(s) 24, 50.

Janssen, Michel (2002*b*). Reconsidering a scientific revolution: the case of Einstein versus Lorentz. *Physics in Perspective (PIP)*, **4**(4), 421–446.
Cited on page(s) 10, 50.

Janssen, Michel (2008, Feb). Drawing the line between kinematics and dynamics in special relativity. PhilSci Archive: #3895. In: [2007] Symposium on Time and Relativity (Minneapolis 25–27 October, 2007).
Cited on page(s) 51.

Janssen, Michel (2009). Drawing the line between kinematics and dynamics in special relativity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **40**(1), 26–52.
Cited on page(s) 16, 20, 24, 33, 48, 49, 50, 51, 55, 69.

Joshi, K. (1983). *Introduction to general topology*. Wiley Eastern Limited.
Cited on page(s) 103, 104.

Kitcher, Philip (1976, Apr). Explanation, conjunction, and unification. *Journal of Philosophy*, **73**(8), 207–212.
Cited on page(s) 42.

Kitcher, Philip (1981, Dec). Explanatory unification. *Philosophy of Science*, **48**(4), 507–531.
Cited on page(s) 42.

Kitcher, Philip (1985, Nov). Two approaches to explanation. *Journal of Philosophy*, **82**(11), 632–639.
Cited on page(s) 42, 43.

Kitcher, Philip (1989). Explanatory unification and the causal structure of the world. See Kitcher and Salmon [1989].
Cited on page(s) 42, 43.

Kitcher, Philip and Salmon, Wesley (ed.) (1989). *Scientific Explanation*, Volume XIII of *Minnesota Studies in the Philosophy of Science Series*. University of Minnesota Press.
Cited on page(s) 223, 229.

Klein, Martin (1967). Thermodynamics in Einstein's thought. *Science*, **157**(3788), 509–516.
Cited on page(s) 15.

Klein, Martin (1982). Some turns of phrase in Einstein's early papers. In *Physics as natural philosophy* (ed. A. Shimony and H. Feshbach), pp. 369–373. MIT Press.
Cited on page(s) 13, 15.

Knox, Eleanor (2013, Oct). Newtonian spacetime structure in light of the equivalence principle. *British Journal for the Philosophy of Science*, **64**(0), 1–18. Advance Access.
Cited on page(s) 94.

Kox, Anne, Klein, Martin, and Schulmann, Robert (ed.) (1997). *The collected papers of Albert Einstein: English translation of selected texts*, Volume 6: The Berlin Years: Writings, 1914–1917. Princeton University Press. Translated by Alfred Engel.
Cited on page(s) 217.

Lange, Marc (2011, Sep). How to explain the Lorentz transformations. Work in progress, sent in personal correspondence.
Cited on page(s) 34.

Lévy-Leblond, Jean-Marc (1963, Jun). Galilei group and nonrelativistic quantum mechanics. *Journal of Mathematical Physics*, **4**(6), 776–788.
Cited on page(s) 123.

Lévy-Leblond, Jean-Marc (1967). Nonrelativistic particles and wave equations. *Communications in Mathematical Physics*, **6**, 286–311.
Cited on page(s) 123.

Lewis, David (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, **61**(4), 343–377.
Cited on page(s) 103.

Lewis, David (1987 [1986]*b*). *Philosophical Papers [electronic resource]*, Volume II. Oxford University Press: Oxford Scholarship Online.
Cited on page(s) 96.

Lewis, David (1993). Mathematics is megethology. *Philosophia Mathematica*, **1**(1), 3–23.

Cited on page(s) 162.

Lewis, David (1994, Oct). Humean supervenience debugged. *Mind*, **103**(412), 473–490.
Cited on page(s) 100, 101.

Lewis, David (2001 [1973]). *Counterfactuals*. Blackwell.
Cited on page(s) 96.

Lewis, David (2001 [1986]*a*). *On the plurality of worlds*. Blackwell.
Cited on page(s) 101, 102, 103, 105.

Linnebo, Øystein (2013). Plural quantification. In *The Stanford Encyclopedia of Philosophy* (Spring 2013 edn) (ed. E. Zalta). The Metaphysics Research Lab: Center for the Study of Language and Information, Stanford University.
Cited on page(s) 162.

Lipschutz, Seymour (1965). *Theory and Problems of General Topology*. Schaum's Outlines Series. Schaum Publishing Co.
Cited on page(s) 109, 110, 184, 185, 188.

Lorentz, Hendrik (1892*a*). *La théorie électromagnétique de Maxwell et son application aux corps mouvants*. Brill. Originally published in *Archives Néerlandaises des sciences exactes et naturelles* 25 (1892), pp. 363–552.
Cited on page(s) 8.

Lorentz, Hendrik (1895). *Versuch einer theorie der electrischen und optischen erscheinungen in bewegten körpern*. Brill.
Cited on page(s) 9, 10.

Lorentz, Hendrik (1916). *The theory of electrons and its applications to the phenomena of light and radiant heat: A course of lectures delivered in Columbia University, New York, in March and April 1906* (2nd edn). Leipzig: B. G. Teubner.
Cited on page(s) 38.

Lorentz, Hendrik (1935 [1900]). Elektromagnetische theorien physikalischer erscheinungen. In *Collected Papers* (ed. M. Nijhoff), Volume VIII, pp. 333–352. The Hague.
Cited on page(s) 46.

Lorentz, Hendrik (1935 [1904]). Ergebnisse und probleme der elektronentheorie. In *Collected Papers* (ed. M. Nijhoff), Volume VIII, pp. 76–124. The Hague.
Cited on page(s) 46.

Lorentz, Hendrik (1937 [1892]*b*). The relative motion of the earth and the ether. In *Collected Papers* (ed. M. Nijhoff), Volume III, pp. 219–223. The Hague.
Cited on page(s) 9.

Malament, David (2012). *Topics in the foundations of general relativity and newtonian gravitation theory*. Chicago Lectures in Physics Series. University of Chicago Press.
Cited on page(s) 94, 170, 183, 186, 191, 192.

Martínez, Alberto (2007). There's no pain in the FitzGerald contraction, is there? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **38**(1), 209–215.
Cited on page(s) 30.

Maudlin, Tim (1993, Jun). Buckets of water and waves of space: Why spacetime is probably a substance. *Philosophy of Science*, **60**(2), 183–203.
Cited on page(s) 65, 95.

Maudlin, Tim (2010, Jun). The geometry of space-time I: Time, topology, and physical geometry. *Aristotelian Society Supplementary Volume*, **84**(1), 63–78.
Cited on page(s) 180, 183, 184.

Maudlin, Tim (2012). *The philosophy of physics: Space and time.* Princeton Foundations of Contemporary Philosophy Series. Princeton University Press.
Cited on page(s) 21, 22, 69.

Maudlin, Tim (2014*a*). *New foundations for physical geometry*, Volume I. Forthcoming by Oxford University Press.
Cited on page(s) 180, 184.

Maudlin, Tim (2014*b*). *New foundations for physical geometry*, Volume II. Forthcoming by Oxford University Press.
Cited on page(s) 180, 181.

Maxwell, James Clerk (1873). *A treatise on electricity and magnetism.* Clarendon Press.
Cited on page(s) 6.

McCormmach, Russell (1970, Winter). H. A. Lorentz and the electromagnetic view of nature. *Isis*, **61**(4), 459–497.
Cited on page(s) 15.

McTaggart, John (1908, Oct). The unreality of time. *Mind*, **17**(68), 457–474.
Cited on page(s) 28.

Michelson, Albert (1881, Aug). The relative motion of the earth and the luminiferous ether. *American Journal of Science*, **22**(128), 120–128.
Cited on page(s) 7.

Michelson, Albert and Morley, Edward (1887, Nov). On the relative motion of the earth and the luminiferous ether. *American Journal of Science*, **34**(203), 333–345.
Cited on page(s) 7.

Miller, Arthur (1981). *Albert Einstein's special theory of relativity: Emergence (1905) and early interpretation (1905–1911).* Addison–Wesley Publishing Company, Inc.
Cited on page(s) 5, 14, 216.

Miller, Willard (1977). *Symmetry and separation of variables*, Volume 4 of *Encyclopedia of Mathematics and its Applications Series.* Addison–Wesley.
Cited on page(s) 133.

Minkowski, Hermann (1909). Raum und zeit. *Physikalische Zeitschrift*, **10**, 104–111. Note: quotations and page numbers taken from English translation [Minkowski 1952].
Cited on page(s) 10, 13, 17, 18, 19, 28, 49, 80, 81, 226.

Minkowski, Hermann (1952). Space and time. In *The principle of relativity: A collection of original memoirs on the special and general theory of relativity* (ed. W. Perrett and G. B. Jeffery), pp. 73–91. Dover. A translation of an address delivered at the 80th assembly of german natural scientists and physicians, at Cologne, 21 September, 1908 [Minkowski 1909].
Cited on page(s) 226.

Misner, Charles, Thorne, Kip, and Wheeler, John (1973). *Gravitation*. W. H. Freeman, San Francisco.
Cited on page(s) 183, 184.

Mulder, Henk and van de Velde-Schlick, Barbara (ed.) (1979*a*). *Moritz Schlick: Philosophical papers*, Volume 11-I (1909-1922) of *Vienna Circle Collection Series*. Reidel Publishing Company.
Cited on page(s) 179, 230.

Mulder, Henk and van de Velde-Schlick, Barbara (ed.) (1979*b*). *Moritz Schlick: Philosophical papers*, Volume 11-II (1925-1936) of *Vienna Circle Collection Series*. Reidel Publishing Company.
Cited on page(s) 179.

Mundy, Brent (1986, Mar). The physical content of Minkowski geometry. *British Journal for the Philosophy of Science*, **37**(1), 25–54.
Cited on page(s) 66.

Naber, Gregory (2012). *The geometry of Minkowski spacetime [electronic resource]* (2nd edn). Number 92 in Applied Mathematical Sciences Series. Springer.
Cited on page(s) 17, 112, 177.

Nanda, Sribatsa (1971, Mar). Topology for Minkowski space. *Journal of Mathematical Physics*, **12**(3), 395–401.
Cited on page(s) 180.

Nanda, Sribatsa (1972, Jan). Weaker versions of Zeeman's conjectures on topologies for Minkowski space. *Journal of Mathematical Physics*, **12**(1), 12–15.
Cited on page(s) 180.

Nerlich, Graham (1976). *The shape of space*. Cambridge University Press.
Cited on page(s) 21.

Nerlich, Graham (1979, Mar). What can geometry explain? *The British Journal for the Philosophy of Science*, **30**(1), 69–83.
Cited on page(s) 21.

Nerlich, Graham (1994*a*). *The shape of space* (2nd edn). Cambridge University Press.
Cited on page(s) 21.

Nerlich, Graham (1994*b*). *What spacetime explains.* Cambridge University Press.
Cited on page(s) 5, 23, 28.

Nerlich, Graham (1998). Time as spacetime. In *Questions of time and tense* (ed. R. L. Poidevin), pp. 119–134. Clarendon Press: Oxford.
Cited on page(s) 28.

Nerlich, Graham (2010). Why spacetime is not a hidden cause: A realist story. In *Space, time, and spacetime: Physical and philosophical implications of Minkowski's unification of space and time* (ed. V. Petkov), Volume 167 of *Fundamental Theories of Physics Series*, Chapter 8, pp. 181–191. Springer.
Cited on page(s) 21, 57.

Nerlich, Graham (2013). *Einstein's genie: Spacetime out of the bottle*, Chapter 5: In SR, why does a rod contract?, pp. 75–98. Minkowski Institute Press.
Cited on page(s) 21, 23, 56, 57, 64.

Newton, Isaac (1704). *Opticks [electronic resource]: Or, a treatise of the reflections, refractions, inflexions and colours of light. Also: Two treatises of the species and magnitude of curvilinear figures.* London.
Cited on page(s) 6.

Newton, Isaac (1729 [1687]). *The mathematical principles of natural philosophy [electronic resource].* London: Printed for Benjamin Motte. Translated by Andrew Motte.
Cited on page(s) 5.

Norton, John (1992). Introduction to the philosophy of space and time. In *Introduction to the philosophy of science: A text by members of the department of the history and philosophy of science of the University of Pittsburgh* (ed. M. Salmon, J. Earman, C. Glymour, J. Lennox, P. Machamer, J. E. McGuire, J. Norton, W. Salmon, and K. Schaffner), Chapter 5, pp. 179–231. Hackett Publishing Company. Reprinted elsewhere by Butterfield *et al.* [1996, pp. . 3–56].
Cited on page(s) 183.

Norton, John (2008). Why constructive relativity fails. *The British Journal for the Philosophy of Science*, **59**(4), 821–834.
Cited on page(s) 59, 60, 61, 87, 90, 106.

Norton, John (2011, Sep). The hole argument. In *The Stanford Encyclopedia of Philosophy* (Fall 2011 edn) (ed. E. Zalta). The Metaphysics Research Lab: Center for the Study of Language and Information, Stanford University.
Cited on page(s) 183.

Oberdan, Thomas (2009). Geometry, convention, and the relativized apriori: The Schlick–Reichenbach correspondence. In *Stationen. Dem philosophen und physiker Moritz Schlick zum 125. Geburtstag* (ed. F. Stadler, H. J. Wendel, and M. von Edwin Glassner), Volume 1 of *Schlick-Studien Series*, pp. 186–211. Springer.
Cited on page(s) 83, 84, 86.

Oberdan, Thomas (2013). Moritz schlick. In *The Stanford Encyclopedia of Philosophy* (Summer 2013 edn) (ed. E. Zalta). The Metaphysics Research Lab: Center for the Study of Language and Information, Stanford University.
Cited on page(s) 75, 83, 84, 85.

Olver, Peter (1986). *Applications of Lie groups to differential equations*. Number 107 in Graduate Texts in Mathematics Series. Springer.
Cited on page(s) 133.

Penrose, Roger and Rindler, Wolfgang (1984). *Spinors and space-time: Two-spinor calculus and relativistic fields*, Volume 1 of *Cambridge Monographs on Mathematical Physics Series*. Cambridge University Press.
Cited on page(s) 169.

Perrett, W. and Jeffery, G. B. (ed.) (1952). *The principle of relativity: A collection of original memoirs on the special and general theory of relativity*. Dover Publications, Inc. An unabridged and unaltered republication of the translation first published by Methuen and Company, Ltd. in 1923.
Cited on page(s) 218.

Petkov, Vesselin (ed.) (2010). *Minkowski spacetime: A hundred years later [electronic resource]* (1st edn), Volume 165 of *Fundamental Theories of Physics Series*. Springer.
Cited on page(s) 215, 219.

Poincaré, Henri (1898, Oct). On the foundations of geometry. *The Monist*, **IX**(1), 1–43.
Cited on page(s) 74.

Poincaré, Henri (1952 [1905]). *Science and Hypothesis*, Chapter V: Experiment and geometry. Dover Publications, Inc.
Cited on page(s) 73, 74.

Pooley, Oliver (2013, Jan). Substantivalist and relationalist approaches to spacetime. In *The Oxford Handbook of Philosophy of Physics* (ed. R. Batterman), Chapter 15, pp. 522–586. Oxford University Press.
Cited on page(s) 16, 30, 32, 62, 64, 65, 82, 94, 95, 106, 108, 157, 159, 170, 174, 175.

Pooley, Oliver (2014). *The reality of spacetime*. In Preparation.
Cited on page(s) 16, 32, 65, 88, 89, 91, 94, 95, 106, 157.

Quine, Willard (1951, Jan). Ontology and ideology. *Philosophical Studies*, **2**(1), 11–15.
Cited on page(s) 1, 65.

Reichenbach, Hans (1928). *Philosophie der raum-zeit-lehre*. Walter de Gruyter & Co. Note: quotations and page numbers taken from translation [Reichenbach 1958].
Cited on page(s) 74, 229.

Reichenbach, Hans (1949). The philosophical significance of the theory of relativity. See Schilpp [1949], Chapter 10, pp. 289–311.
Cited on page(s) 74.

Reichenbach, Hans (1958). *The philosophy of space and time*. Dover Publications. Translation of *Philosophie der Raum-Zeit-Lehre* [Reichenbach 1928].
Cited on page(s) 228.

Reichenbach, Hans (1969). *Axiomatization of the theory of relativity*. University of California Press. First published in German under the title *Axiomatik der relativistischen Raum-Zeit-Lehre*. Translated and edited by Maria Reichenbach.
Cited on page(s) 19.

Robb, Alfred (1914). *A theory of time and space*. Cambridge University Press.
Cited on page(s) 111.

Robb, Alfred (1936). *The geometry of time and space*. Cambridge University Press.
Cited on page(s) 111.

Rostock, Universität (2013). Moritz-Schlick-Forschungsstelle. URL: <http://www.moritz-schlick.de>.
Cited on page(s) 179.

Salmon, Wesley (1971). *Statistical explanation & statistical relevance*. University of Pittsburgh Press. With contributions by Richard Jeffrey and James Greeno.
Cited on page(s) 41.

Salmon, Wesley (1984). *Scientific explanation and the causal structure of the world*. Princeton University Press.
Cited on page(s) 42.

Salmon, Wesley (1989). Four decades of scientific explanation. See Kitcher and Salmon [1989], pp. 3–219. Reprinted elsewhere by Salmon [1990].
Cited on page(s) 40, 229.

Salmon, Wesley (1990). *Four decades of scientific explanation*. University of Minnesota Press, Minneapolis. A reprinting of the original publication [Salmon 1989].
Cited on page(s) 229.

Salmon, Wesley (1992). Scientific explanation. In *Introduction to the philosophy of science: A text by members of the department of the history and philosophy of science of the University of Pittsburgh*, Chapter 1, pp. 7–41. Prentice Hall.
Cited on page(s) 41.

Salmon, Wesley (1994, Jun). Causality without counterfactuals. *Philosophy of Science*, **61**(2), 279–312.
Cited on page(s) 42.

Salmon, Wesley (1997, Sep). Causality and explanation: A reply to two critiques. *Philosophy of Science*, **64**(3), 461–477.
Cited on page(s) 42.

Salmon, Wesley (1998). *Causality and Explanation*. Oxford University Press.
Cited on page(s) 40, 41, 42, 43.

Salmon, Wesley (2011). *Causality and Explanation [electronic resource]*. Oxford University Press: Oxford Scholarship Online.
Cited on page(s) 43, 44.

Saunders, Simon (2003). Physics and Leibniz's principles. In *Symmetries in physics: Philosophical reflections* (ed. K. Brading and E. Castellani), Chapter 16, pp. 289–308. Cambridge University Press.
Cited on page(s) 89.

Saunders, Simon (2006, Jan). Are quantum particles objects? *Analysis*, **66**(1), 52–63.
Cited on page(s) 89.

Schechter, Eric (1997). *Handbook of analysis and its foundations*. Academic Press.
Cited on page(s) 188, 189, 200.

Schechter, Eric (1998). *Handbook of analysis and its foundations [electronic resource]* (CD-ROM edn). Academic Press.
Cited on page(s) 189.

Schilpp, Paul (ed.) (1949). *Albert Einstein: Philosopher-scientist* (1 edn), Volume VII of *The Library of Living Philosophers Series*. Cambridge University Press.
Cited on page(s) 217, 228.

Schlick, Moritz (1915). Die philosophische bedeutung des relativitätsprinzips. *Zeitschrift für Philosophie und philosophische Kritik*, **159**, 129–175. An English translation is also available [Schlick 1979].
Cited on page(s) 83, 230.

Schlick, Moritz (1949). *Philosophy of nature*. Philosophical Library, Inc. Translated by Amethe von Zeppelin.
Cited on page(s) 85.

Schlick, Moritz (1974 [1925]). *General Theory of Knowledge*. Library of Exact Philosophy Series. Springer-Verlag. Translated by Albert Blumberg, from the 2nd German Edition of *Allgemeine Erkenntnislehre*, Berlin 1925.
Cited on page(s) 83, 85, 86.

Schlick, Moritz (1979). The philosophical significance of the principle of relativity. See Mulder and van de Velde-Schlick [1979*a*], Chapter 7, pp. 153–189. Translated by Peter Heath from the original [Schlick 1915].
Cited on page(s) 230.

Schlick, Moritz (1979 [1922]). Space and time in contemporary physics: An introduction to the theory of relativity and gravitation. See Mulder and van de Velde-Schlick [1979*a*], Chapter 9, pp. 207–269. Translated by Henry Brose and Peter Heath from *Raum und zeit in der gegenwärtigen physik*, 4th edition, Berlin 1922.
Cited on page(s) 83, 84, 85, 86.

Seelig, Carl (ed.) (1954). *Ideas and opinions*. Crown Publishers, Inc.
Cited on page(s) 217, 218.

Sklar, Lawrence (1974). *Space, time, and spacetime*. University of California Press.
Cited on page(s) 183.

Skow, Bradford (2006, May). Physical relativity [book review]: Space-time structure from a dynamical perspective. In *Notre Dame Philosophical Reviews*. University of Notre Dame. Available online: <http://ndpr.nd.edu/review.cfm?id=6603>.
Cited on page(s) 30, 75.

Skow, Bradford (2013). Are there non-causal explanations (of particular events)? *British Journal for the Philosophy of Science*, **0**(0), 1–23. Advance Access: May 31, 2013.
Cited on page(s) 21.

Stachel, John (ed.) (1989*a*). *The collected papers of Albert Einstein: English translation of selected texts*, Volume 2: The Swiss Years: Writings, 1900–1909. Princeton University Press. Translated by Anna Beck.
Cited on page(s) 216.

Stachel, John (1989*b*). The rigidly rotating disk as the "missing link" in the history of general relativity. In *Einstein and the development of general relativity: Based on the proceedings of the 1986 Osgood Hill conference, North Andover, Massachusetts, 8–11 May 1986* (ed. D. Howard and J. Stachel), Volume 1 of *Einstein Studies Series*, pp. 48–62. Birkhäuser. Originally published as "Einstein and the Rigidly Rotating Disk" in *General Relativity and Gravitation* (1980).
Cited on page(s) 73.

Stachel, John (ed.) (1998). *Einstein's miraculous year*. Princeton University Press.
Cited on page(s) 216.

Stachel, John (2002). *Einstein from 'B' to 'Z'*, Volume 9 of *Einstein Studies Series*. Birkhäuser.
Cited on page(s) 5.

Stachel, John and Janssen, Michel (2004). The optics and electrodynamics of moving bodies. To appear in Italian in the encyclopedic series, *Storia Della Scienza*, and also in English in Stachel's forthcoming book, *Going Critical*, Kluwer. Preprint available online: <http://www.mpiwg-berlin.mpg.de/preprints/p265.pdf>.
Cited on page(s) 5, 7, 10.

Stark, Glenn (2013). *Light*. Encyclopædia Brittanica Inc.
Cited on page(s) 6.

Sypel, Roland and Brown, Harvey (1992). When is a physical theory relativistic? In *PSA 1992: Proceedings of the 1992 biennial meeting of the Philosophy of Science Association* (ed. D. Hull, M. Forbes, and O. Kathleen), Volume 1, pp. 507–514 (Part VIII). Philosophy of Science Association.
Cited on page(s) 30.

van Camp, Wesley (2011*a*, Dec). On kinematic versus dynamic approaches to Special Relativity. *Philosophy of Science*, **78**(5), 1097–1107.
Cited on page(s) 34, 45, 51.

van Camp, Wesley (2011*b*, Feb). Principle theories, constructive theories, and explanation in modern physics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **42**(1), 23–31.
Cited on page(s) 45.

van Fraassen, Bas (1970). *An introduction to the philosophy of time and space* (Morningside edn). Studies in Philosophy Series. Columbia University Press.
Cited on page(s) 98, 164.

Wachter, Armin (2011). *Relativistic quantum mechanics [electronic resource]*. Theoretical and Mathematical Physics Series. Springer.
Cited on page(s) 122, 125.

Wald, Robert (1984). *General relativity*. The University of Chicago Press.
Cited on page(s) 183, 184, 185.

Weatherson, Brian (2010). David Lewis. In *The Stanford Encyclopedia of Philosophy* (Summer 2010 edn) (ed. E. Zalta). The Metaphysics Research Lab: Center for the Study of Language and Information, Stanford University.
Cited on page(s) 162.

Winnie, John (1977). The causal theory of space-time. In *Foundations of Space-Time Theories* (ed. J. Earman, C. Glymour, and J. Stachel), Volume VIII of *Minnesota Studies in the Philosophy of Science Series*, Chapter 7, pp. 134–205. University of Minnesota Press.
Cited on page(s) 19, 110, 111, 177.

Woodhouse, Nick (2003). *Special relativity*. Springer.
Cited on page(s) 164.

Woodward, James (2010). Scientific explanation. In *The Stanford Encyclopedia of Philosophy* (Spring 2010 edn) (ed. E. Zalta). The Metaphysics Research Lab: Center for the Study of Language and Information, Stanford University.
Cited on page(s) 40, 41, 43.

Young, Thomas (1804). The Bakerian lecture: Experiments and calculations relative to physical optics. *Philosophical Transactions of the Royal Society of London*, **94**, 1–16.
Cited on page(s) 6.

Zahar, Elie (1989). *Einstein's revolution: A study in heuristic*. Open Court.
Cited on page(s) 5, 9, 10.

Zeeman, Erik (1964, Apr). Causality implies the Lorentz group. *Journal of Mathematical Physics*, **5**(4), 490–493.
Cited on page(s) 111.

Zeeman, Erik (1967, Apr). The topology of Minkowski space. *Topology*, **6**(2), 161–170.
Cited on page(s) 112, 176.