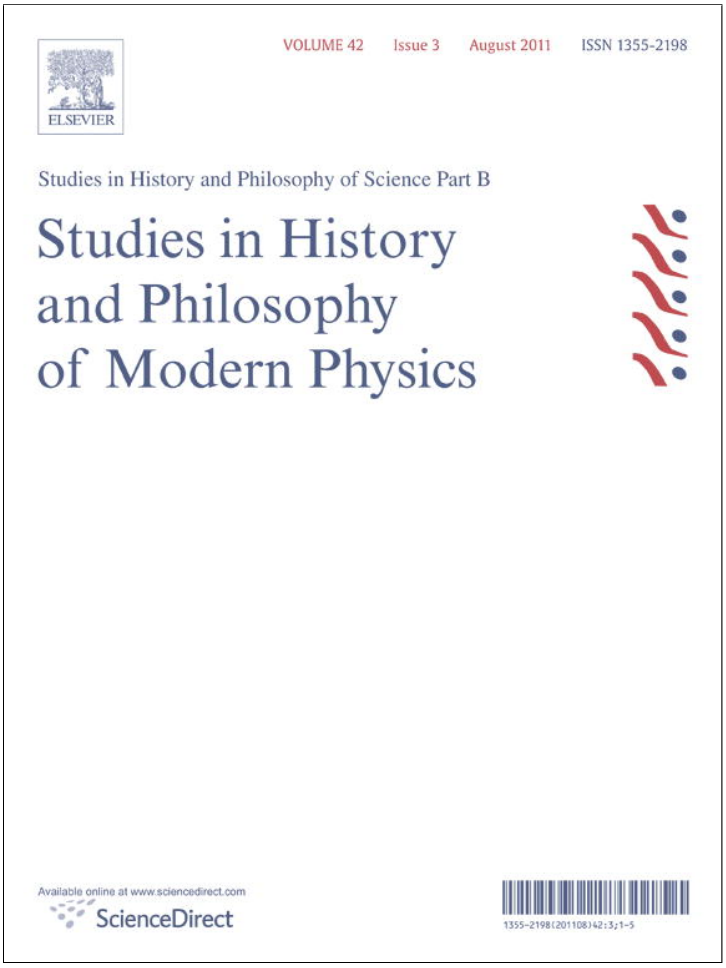
VOLUME 42 Issue 3 August 2011 ISSN 1355-2198
ELSEVIER
Studies in History and Philosophy of Science Part B
Studies in History and Philosophy of Modern Physics
Available online at www.sciencedirect.com
ScienceDirect

Studies in History and Philosophy of Modern Physics 42 (2011) 176–183

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

Principle or constructive relativity

Mathias Frisch

University of Maryland, Department of Philosophy, College Park, MD 20742, USA

ARTICLE INFO

Article history:
Received 24 June 2010
Received in revised form
27 March 2011
Accepted 13 April 2011

Keywords:
Relativity
Einstein
Lorentz
Principle theories
Constructive theories
Kinematics
Dynamics

ABSTRACT

Appealing to Albert Einstein's distinction between principle and constructive theories, Harvey Brown has argued for an interpretation of the theory of relativity as a dynamic and constructive theory. Brown's view has been challenged by Michel Janssen and in this paper I investigate their dispute. I argue that their disagreement appears larger than it actually is due to the two frameworks used by Brown and Janssen to express their respective views: Brown's appeal to Einstein's principle–constructive distinction and Janssen's framing of the disagreement as one over the question whether relativity provides a kinematic or a dynamic constraint. I appeal to a distinction between types of theories drawn by H. A. Lorentz two decades before Einstein's distinction to argue that Einstein's distinction represents a false dichotomy. I argue further that the disagreement concerning the kinematics–dynamics distinction is a disagreement about labels but not about substance. There remains a genuine disagreement over the explanatory role of spacetime geometry and here I agree with Brown arguing that Janssen sees a pressing need for an explanation of Lorentz invariance where no further explanation is needed.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Modern Physics*

1. Introduction

It is generally believed that Einstein's special theory of relativity has changed our understanding of space and time and of the motion of objects in two important ways. First, we no longer think, as adherents of the classical electromagnetic world-picture at the turn of the twentieth century did, that there is a privileged class of inertial frames, the 'ether rest frames'; second, it seems to be widely believed that relativistic phenomena, such as length contraction, do not require a detailed explanation in terms of electromagnetic forces or quantum mechanical interactions—that is, a dynamical explanation—but that they are artifacts of the relative state of motion of the frames used for representation and are simply a consequence of the geometry of spacetime. This orthodox view has been challenged by Brown and Pooley (Brown, 2005; Brown & Pooley, 2006), who agree with the first part of the orthodoxy but not with the second. According to what Brown takes to be the orthodox view, the structure of spacetime—that is, the fact that spacetime is Minkowskian—explains the fact that the laws of our theories are Lorentz-invariant and thereby accounts for the universal behavior of rods and clocks. But Brown argues that this view has the arrow of explanation backward and that relativistic phenomena ultimately require a dynamical explanation. According to Brown, it is a brute fact, which itself requires no further explanation, that the laws are Lorentz-invariant and it is this fact, which explains that Minkowski spacetime is the proper arena to represent non-gravitational physical phenomena.

Brown's view has recently been forcefully criticized by Janssen (2009) and in this paper I want to investigate their dispute. I argue that there is less disagreement between the two positions than may appear initially and, hence, that Brown's view presents less of a departure from orthodoxy than it may seem. In Section 2 I provide a first summary of Brown's view and Janssen's response. In Section 3 I argue that at least one source of their apparent disagreement is the role played by Einstein's well-known distinction between principle and constructive theories in the way in which especially Brown but also Janssen frame their respective views. Thus, in a sense I agree with Janssen, who believes that "the principle–constructive distinction is a red herring in the end" (Janssen, 2009, p. 38). I appeal to an earlier, related distinction due to H. A. Lorentz, involving a notion of general principle broader than Einstein's, to argue that Einstein's distinction presents a false dichotomy. Once we locate Brown's and Janssen's respective positions within Lorentz's framework, I argue in Section 4, a significant part of disagreement between their respective views disappears.

E-mail address: mfrisch@umd.edu

1355-2198/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.shpsb.2011.04.004

In Section 5 I argue that the fact that Brown maintains that the theory of relativity provides dynamical explanations whereas Janssen takes it to provide a kinematical constraint also does not constitute a genuine disagreement. The true disagreement between Janssen and Brown concerns the question whether Lorentz invariance ought to be explained in terms of the geometry of spacetime or whether the arrow of explanation is the other way around. While both Brown and Janssen themselves have also characterized their disagreement in these terms, this issue gets obscured by Brown's use of Einstein's distinction and Janssen's characterization of the disagreement in terms of the kinematics–dynamics distinction—or so I shall argue. As far as the remaining disagreement between Brown and Janssen is concerned, I agree with Brown. In particular, Janssen sees a pressing need for an explanation of the principle of Lorentz invariance, whereas, as I will argue in Section 6, there is no such need. I focus exclusively on special relativity in this paper.

2. Physical relativity: first pass

One of Brown's targets is the view that spacetime substantivalism plays an important role in explaining relativistic effects. The substantivalist takes spacetime to be an entity in its own right. Once we know that objects 'live' in Minkowski spacetime, and satisfy the constraints of Minkowski geometry, there is a simple well-known geometric construction that allows us to derive length contraction and time dilation. Thus, length contraction and time dilation appear to be purely geometric effects, which are a straightforward consequence of the structure of Minkowski spacetime. But Brown argues that merely appealing to the structure of spacetime does not answer the question as to *why* objects obey the constraints of Minkowski geometry. For him the question as to how rods and clocks might be able to know in what spacetime they are immersed is the "mystery of mysteries" (Brown, 2005, p. 143). This point is echoed in Brown's discussion of geometric explanations of inertial effects. In positing that force-free objects simply follow the geodesics of Minkowski spacetime, determined by its affine structure, Brown argues, the substantivalist needs to assume that objects have "space–time feelers" (Brown, 2005, p. 24) that are somehow able to sense "the ruts and grooves" of spacetime. He concludes that spacetime structure can play no explanatory role and merely serves "as codification of certain key aspects of the behavior of particles (and/or fields)" (Brown, 2005, p. 25).

By contrast, what can explain the behavior of rods and clocks are "the details of the bodies' microphysical constitution" (Brown & Pooley, 2006, p. 76): the explanation of Lorentz contraction "is ultimately to be sought in terms of the dynamics of the microstructure of the contracting rod" (Brown, 2005, p. 133). A moving rod contracts, that is, "*because of how it is made up and not because of the structure of its spatio-temporal environment*" (Brown, 2005, p. 8, italics in original). Thus, Brown argues for the priority of the dynamical laws over spacetime structure. A distinction that plays a central role in the way in which Brown frames his account is Einstein's well-known distinction between principle and constructive theories. Put in terms of this distinction Brown's thesis becomes the claim that the behavior of rods and clocks calls for a constructive explanation, which provides a dynamical model of the objects' detailed microscopic structure.

Against Brown, Janssen defends the view that the behavior of rods and clocks ultimately is explained by the geometric structure of spacetime. Janssen agrees with Brown in his rejection of spacetime substantivalism but argues that even within a relationalist framework, which takes spatio-temporal relations to have no existence independent of the physical objects that instantiate them, the explanatory arrow is from the structure of Minkowski spacetime to the Lorentz invariance of the dynamical laws. As Janssen puts it: "I argue that the space–time symmetries are the explanans and that the Lorentz invariance of the various laws is the explanandum; Brown argues that it is the other way around" (Janssen, 2009, p. 26).[1]

Central to Janssen's argument against Brown are extensive case studies of several phenomena for which pre-relativistic explanations in terms of detailed micro-dynamical models were eventually replaced by relativistic explanations that show that the phenomena in question are independent of the specifics of the dynamics. That is, contrary to what Brown seems to suggest, the details of the bodies' microphysical constitution play no role in the ultimate explanation of these phenomena. According to Janssen's terminology, this means that the phenomena are "kinematical in the broad sense". Janssen then argues that the fact that the phenomena are kinematical in the broad sense is explained by the fact that they are also what he calls "kinematical in the narrow sense": the phenomena are instances of standard spatio-temporal behavior. Thus, he concludes that instead of a need to appeal to distinct detailed dynamical models for different relativistic phenomena "the statement that space–time is Minkowskian explains all in one fell swoop" (Janssen, 2009, p. 49).

On first sight, then, there seems to be a rather stark disagreement between Brown's and Janssen's views. Brown, as many seem to read him, appears to believe that only an account of the particular forces pushing and pulling the microscopic constituents of a rod can explain length contraction; Janssen by contrast, maintains not only that the explanatory advantage of Einstein's theory over Lorentz's consists precisely in the fact that the special theory of relativity teaches us that certain phenomena are independent of the detailed assumptions about the dynamics governing a particular system, but he insists, moreover, that it is the structure of Minkowski spacetime that explains this independence. Brown argues that the behavior of rods and clocks ultimately calls for a dynamical explanation; while Janssen argues that the explanation is kinematic. While both agree that the correct or best explanation of length contraction is "constructive", in terms of Einstein's terminology, for Janssen the correct constructive explanation appeals to the geometry of spacetime rather than to the micro-dynamics of rods.

Curiously both Brown and Janssen appeal to Wolfgang Pauli in support of their respective views. Brown approvingly quotes (at least three times in his book!) Pauli's claim that "the contraction of a measuring rod is not an elementary but a very complicated process" (Pauli, 1921, p. 15). Janssen cites Pauli as saying that "it constituted a definite progress that Lorentz's law of the variability of mass could be derived from the theory of relativity without making any specific assumptions on the electron shape or charge distribution. Also nothing need be assumed about the nature of the mass" (cited in Janssen, 2009, pp. 38 and 39). While the first quote appears to suggest that a proper understanding of length contraction needs to invoke the detailed and complicated dynamics governing the rod, the second quote seems to imply that it is precisely such a detailed model that relativity theory has shown to be superfluous in accounting for the so-called 'relativistic phenomena.' Yet, contrary to first impressions, it is not clear that there really is a tension between the two quotes and the fact

[1] John Norton has argued that Brown's spacetime relationalism fails since it needs to antecedently presumes essential commitments of a realist conception of spacetime (Norton, 2008, Sections 1–5). But Norton's relationalist not only denies the existence of a four-dimensional substantival spacetime that exists independently of matter but also that matter has no basic spatio-temporal properties. A relationalist who merely denies the former claim can escape Norton's conclusion.

that both Brown and Janssen cite Pauli's review article on relativity in support of their respective views can serve as a first clue that their disagreement might be less than what initial appearances suggest. Since I believe that Einstein's principle vs. constructive theory distinction obscures rather than helps to elucidate where Brown and Janssen agree and where they disagree, I want to begin my reconstruction with a discussion of this distinction.

3. Principles and mechanisms

Albert Einstein, in a letter to the London *Times*, famously introduced a distinction between what he called "constructive theories" and "principle theories" (Einstein, 1919 [1954], p. 228). According to Einstein, we can distinguish constructive theories that treat phenomena as complex and build them up out of "the materials of a relatively simple formal scheme" from principle theories that rely on empirically discovered, general characteristics of natural processes. As it is usually understood, Einstein's distinction is a distinction between theories that describe the directly observable macroscopic behavior of a system with the help of phenomenological principles, which are elevated to the status of postulates, and theories that describe a system's behavior by proposing a model of its detailed microscopic constitution. This interpretation fits well with the examples Einstein gives of the two kinds of theories: the kinetic theory of gases, on one hand, which derives the behavior of gases from a microscopic model, and thermodynamics, on the other hand, which is based on phenomenological principles such as the principle that there can be no perpetual motion.

Einstein says that both kinds of theory have advantages—constructive theories are complete, adaptable, and clear, while the advantage of principle theories are "logical perfection and secureness of foundations"—yet he is unequivocal in his preference for constructive theories: "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 process in question" (Einstein, 1919). That is, according to Einstein, we simply do not understand a natural process unless we are in possession of a constructive theory of that process. Brown and Pooley follow Einstein in their assessment of the relative merits of the two kinds of theory. They, too, believe not only that constructive theories are explanatorily superior to principle theories but also that "principle theories fail to be explanatory" (Brown & Pooley, 2006, p. 74). Thus, principle theories are "explanatorily inferior," according to Brown and Pooley in quite a dramatic fashion: they simply "fail to provide any sort of explanation" (Brown & Pooley, 2006)! This is a verdict with which Janssen agrees as well.[2]

In his letter to the *Times* Einstein maintains that special relativity is a principle-theory, based on the relativity postulate and the light postulate. The behavior of rods and clocks is derived in the theory by showing that in any world in which the relativity postulate and the light postulate holds all objects, independently of their microphysical constitution, must necessarily exhibit the phenomena of length construction and time dilation. Now, at first sight it might seem rather surprising that Einstein, in a letter devoted to outlining his theory of relativity, would deny that this theory could yield understanding. Brown proposes the following interpretation of this fact. In 1905 no satisfactory constructive theory that could account for the stability of matter was available and a principle-theory approach that simply postulated stable macroscopic measuring rods and clocks was the only option. But this approach was merely preliminary for Einstein, who, according to Brown, believed that a full understanding of these phenomena had to await the development of a satisfactory constructive micro-theory of matter. Einstein himself, of course, never was happy with the micro-theory that was eventually to be developed—that is, with quantum mechanics—but according to Brown it is precisely that theory, perhaps in the form of a quantum field theory, that provides us with the ultimate explanation of length contraction "in terms of the dynamics of the microstructure of the contracting rod" (Brown & Pooley, 2006, p. 77).

That is, Brown maintains that Einstein's 1905 principle-theory approach played an important historical role—it provided logically secure foundations for the treatment of a class of phenomena for which no promising constructive approach was available at the time—but a genuine understanding of the behavior of rods and clocks had to await the development of a constructive theory that provided a microscopic dynamical model for rods and clocks. The principle-theory approach to relativity shows that rods and clocks must behave in quite peculiar ways for the two postulates of the theory to be true together. But this, Brown and Pooley insist, does not constitute an explanation of the behavior of rods and clocks. Rather, explaining length contraction "involves solving the dynamics governing the structure of the complex material body that undergoes contraction" (Brown & Pooley, 2006, p. 79). According to Brown, then, it is a mistake to think of special relativity only, or even primarily, as a principle theory. Rather, the proper way to understand relativity theory *post*-1905 is as a constructive theory.

As we have seen, Brown and Janssen distinguish a third kind of approach to relativity theory, in addition to Einstein's 1905 principle theory and the constructive approach that Brown himself favors: a geometric approach explaining the behavior of rods and clocks by appealing to the structure of Minkowski spacetime. According to Janssen, the geometric interpretation of the theory also provides a constructive explanation of the relativistic phenomena with Minkowski spacetime providing the constructive model. A geometric constructive explanation of length contraction, for example, proceeds "by showing that two observers who are in relative motion to one another and therefore use different sets of space–time axes disagree about which cross-sections of the 'world-tube' of a physical system give the length of the system" (Balashov & Janssen, 2003, p. 331).

In response, Brown and Pooley maintain that there are two distinct spacetime interpretations of relativity, neither of which, however, provides an adequate explanation of the behavior of rods and clocks. First, one might posit Minkowski spacetime as an ontologically autonomous entity and maintain that the behavior of rods and clocks is constrained to reflect the geometric structure of the spacetime they inhabit. Brown concedes that this approach is genuinely constructive, since it posits Minkowski spacetime as reality 'behind the phenomena'. Yet he maintains, as we have already seen, that it fails to explain the behavior of rods and clocks, since it has to posit a mysterious influence of spacetime on objects: "How," he asks, "are rods and clocks supposed to know which space–time they are immersed in?" (Brown, 2005, p. 143). Second, one might try to appeal to the constraints of Minkowski spacetime within a relationalist framework, as proposed by Janssen. Brown and Pooley reject this account on the grounds that it is not truly constructive, since the geometric features in question are phenomenological and do not directly "concern the details of the bodies' microphysical constitution" (Brown & Pooley, 2006, p. 76). Hence, the geometric features have a status akin to the postulates of a principle theory and, thus, are

[2] See Janssen (2009, p. 38, fn. 27): "Brown & Pooley (2006, pp. 74 and 75) correctly point out that, contrary to what Balashov and I suggested, principle theories are not explanatory. Explanations are about the reality behind the phenomena (be it about their causes or about their nature)."

not explanatory: "They are about aspects of their (fairly) directly observable macroscopic behavior. And this reflection prompts an obvious question: *why* do these objects obey the constraints of Minkowski geometry?" (Brown & Pooley, 2006, p. 76).

While I will disagree immediately below with Brown and Pooley as far as their claims about explanatoriness are concerned, I agree with their worry that it is not clear how well Janssen's classification of a spacetime account as constructive sits with Einstein's own characterization of constructive theories as building "up a picture of the more complex phenomena out of the materials of the relatively simple formal scheme from which they start out." In Einstein's example—the kinetic theory of gases—mechanical, thermal and diffusional processes are reduced to the movement of molecules. That is, the theory shows that the macroscopic processes within its scope are the result of extremely complex combinations of simple microscopic building blocks. 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.

According to Einstein's framework, scientific theories fall into one of two categories: purely phenomenological principle theories, on one hand, and constructive theories that provide a detailed microscopic model of the reality behind the phenomena, on the other. Since only theories of the latter kind can be explanatory, relativity theory would have to be a constructive theory, if it were to be explanatory. But as presented so far, Brown's account seems deeply puzzling. Has the lesson of the theory of relativity not precisely been that a wide class of phenomena that might have appeared to depend on the microstructure of objects is independent of the details of their constitution? This, of course, is the point that Janssen stresses forcefully in his case studies in 2009. But while I agree with Janssen on this point, it seems to me that this criticism of Brown is misdirected. To be sure, given the way Brown attempts to locate his interpretation of the theory of relativity within the framework provided by Einstein's principle–constructive distinction, it is difficult not to read him as advocating that a proper explanation of relativistic phenomena would have to involve solving the detailed microscopic equations of motion. But, I want to suggest that part of the fault lies with Einstein's distinction, construed as exhaustive, and that there is a better framework available to capture Brown's view—a view which, I want to claim, is compatible with at least part of the lesson Janssen wishes to draw from his case studies.

Even though the distinction between principle- and constructive theories is usually traced to Einstein, a related distinction had almost two decades earlier been drawn by Lorentz (see Frisch, 2005). Lorentz (1900) distinguishes theories that begin by postulating "general principles" or "general laws" from theories that postulate a "mechanism of the phenomena". General principles, according to him, have the advantage of being versatile and applying to a wide variety of phenomena, since they abstract from and are independent of "the inner constitution of bodies." Mechanism theories, by contrast, provide us "with flawed, yet lively representations" of the "connections between and the nature of things." On first sight this distinction might strike one as being equivalent to Einstein's later distinction between phenomenological and constructive micro-theories. Similar to Einstein, Lorentz takes principles to express "generalized experiences" and Lorentz's mechanisms seem to be the same thing as Einstein's constructive models. But Lorentz's notion of principle is broader than Einstein's (at least as it is generally interpreted), for as an example of a general principle Lorentz cites not only the purely phenomenological second law of thermodynamics but also the principle of energy conservation. Even if it was correct that we arrived at this principle as the result of generalizing our experiences, the principle clearly is not purely phenomenological; rather it has come to function as a general, albeit perhaps defeasible, constraint on all physical theorizing. Thus, there is an important category of physical principles that Lorentz at least implicitly recognizes, but Einstein's distinction between purely phenomenological principles and detailed constructive models of the phenomena leaves out: general constraints that guide theory construction and theory choice and that we take every physical theory, or perhaps every theory within a certain domain, to satisfy. Indeed, it seems helpful to distinguish explicitly between the two kinds of principle theories Lorentz allows, and make the following tripartite distinction:

(i) mechanism or constructive theories, such as the kinetic theory of gases;
(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.

Once we recognize the existence of general constraints as an integral part of physical theorizing, Einstein and Brown's suggestion that only a detailed constructive account of a phenomenon can offer an explanation of that phenomenon becomes suspect, while Lorentz's discussion seems closer to the mark. Lorentz preferred theories of mechanism, just as Einstein preferred constructive theories, but while Einstein maintained that *only* constructive theories can yield understanding, Lorentz stressed that principle theories, too, can yield "insight" and prefaced his discussion of principle and mechanism theories with a commitment to an explanatory pluralism according to which "there are multiple ways by which we try to understand natural phenomena [. . .] Individual characteristics and inclinations determine the choice for each scientist".[3] That is, Lorentz unlike Einstein believed that both (or perhaps even all three) kinds of theories or principles could yield understanding.

I wish to side with Lorentz's more ecumenical view on scientific explanation against the view expressed by Einstein, Brown, and Janssen. To explain a phenomenon, I want to submit, is to embed the phenomenon into a pattern of functional dependencies—in Jim Woodward's terminology it is to answer "what-if-things-had been-different-questions"—and phenomenological principles can provide us with answers to such questions just as general principles or constructive theories can. The second law of thermodynamics can provide insight into why milk mixes in coffee; the Bohr model of the atom offers a constructive explanation of the spectral lines of hydrogen; the principle of energy–momentum conservation can explain why a bouncing ball does not bounce higher than the height from which it was released. There may be important differences in the depth of understanding that various kinds of explanation can provide, but it seems to me to be a mistake to deny that principle or purely phenomenological theories can provide any explanations at all. Rather I take it that explanation is a highly context-dependent notion and that that there may even be contexts in which a phenomenological account can provide the best explanation, just

[3] I take it that am here assuming that the notions of explanation, understanding, and of gaining insight are closely related: to gain insight into a phenomenon is to increase one's understanding of the phenomenon and explanations provide understanding or insight.

as there are others in which a constructive account is called for or where an appeal to a general principle provides the simplest and best explanation. Some explanations may explain by getting at the reality behind a phenomenon or by positing a common origin for a range of different phenomena, as Janssen maintains, but such "common origin explanation" do not exhaust the notion of explanation. Rather it seems to me that Lorentz got it exactly right, when he insisted on the plurality of kinds of explanatory accounts.[4]

Arguably such an explanatory pluralism is also the upshot of J. S. Bell's "Lorentzian pedagogy": As Brown (2005, pp. 124 and 126) explains, Bell argued that it is pedagogically useful to derive the longitudinal contraction of the orbit of a classical electron that is orbiting a moving nucleus dynamically from the Maxwell equations, rather than simply appealing directly to the Lorentz transformations. Solving the dynamical equations explicitly can give us additional insight into relativistic phenomena. Significantly, however, Bell stresses the understanding that can be gained from *both* kinds of derivations: those that rely simply on a general feature of the dynamics—that is, its Lorentz invariance—and those that show how this feature is implemented in a particular case.

I am not claiming that the roles played by the different types of theory or principle cannot overlap. General principles, such as the principle of energy conservation, tend to have broader scope than the other two types of theories, but a constructive 'theory of everything' would obviously also be a theory of universal scope. Also, while I have characterized general principles as providing (defeasible) constraints on all levels of theorizing, constructive theories and phenomenological principles can, of course, also be thought of as providing constraints on future theorizing: If we are committed to a phenomenological theory *P*, then any future constructive theory for the phenomena in *P*'s domain has to be able to 'save' the phenomenological regularities posited by *P*. Perhaps existing successful constructive theories can be seen as functioning as heuristic guides for the development of future constructive theories—if a certain mechanism is successful in accounting for the phenomena in one domain, we're inclined to try the same kind of mechanism elsewhere—and constructive theories thereby can also provide at least a weak constraint on theory development.

My main aim in the present section has been to argue that the distinction between purely phenomenological putatively non-explanatory theories, on one hand, and constructive micro-theories, that aim to explain by positing detailed microphysical mechanisms, on the other, presents a false dichotomy. If we do not think that special relativity provides us with phenomenological principles of arguably limited explanatory depth, similar to the second law of thermodynamics, then we are not thereby forced to conceive of the theory as a constructive theory that, akin to the kinetic theory of gases, relies on a detailed microscopic dynamical story to account for the behavior of rods and clocks.

4. The 'big principle'

What, then, is the theory of special relativity, in terms of Lorentz's framework? It is a *general principle theory* with the principle of Lorentz invariance as its core—a principle, which is not merely phenomenological but which, like the principle of energy conservation, functions as a general constraint on all levels of theorizing, from phenomenological macro-theories to constructive micro-theories of matter. On my reading, then, there is a deep disanalogy between the principle of Lorentz invariance and thermodynamic principles. The latter are purely phenomenological, while the former is not. The theory is not committed to accepting measurement rods and clocks as not further analyzable fundamentals into the theory, contrary to what Einstein's 1905 paper might be taken to suggest. Rods and clocks do have a microstructure, and providing a model of this microstructure can be explanatory, for example of the stability of macroscopic objects. Yet this is not part of the theory of relativity itself, and relativity is no more a constructive theory than it is purely phenomenological. In particular, details of the micro-theory are not needed to account for the relativistic behavior of stable macroscopic objects, which is explained solely by appealing to the principle of Lorentz invariance.

Brown himself, in fact, seems to have recognized the limited value of Einstein's distinction and struggles in his book to locate his answer to the question "what is special relativity?" within Einstein's framework, despite his initial commitment to a constructive account. Brown's ultimate view, it seems to me, is exactly the view I sketched here: the content of special relativity is given by the 'big principle' of Lorentz invariance (see Brown, 2005, pp. 146 and 147). But his problem then is how to think of this principle, which neither has the character of a purely phenomenological principle nor does it itself provide a detailed constructive model. Thus Brown concludes that "at its most fundamental, SR is a theory that lies somewhere in between a pure principle theory (like thermodynamics, or Einstein's 1905 version of SR) and a fully constructive theory (like statistical mechanics)" (Brown, 2005, p. 147). In light of his initial insistence on the superiority of constructive accounts and his earlier discussion of explanation, this verdict is rather surprising, especially since it commits him to the claim that at its most fundamental relativity lies somewhere in between an explanatory theory and one that simply fails to explain!

Finally, thinking of relativity theory as a *general* principle theory allows us to make sense of the two quotations from Pauli above. According to Janssen's reading of the quotations, Pauli's claim that the contraction of a rod is a complicated process confuses the problem of the stability of a rod at rest—which indeed depends on the details of the dynamics governing the rod's microscopic constituents—with the contraction of a moving rod, which follows simply from the fact that the dynamics needed to solve the first problem is Lorentz-invariant. But there is also a more charitable interpretation of what Pauli says. Pauli might be responding to the operationalist streak in Einstein's early expositions of the theory of relativity that seem to presuppose rods and clocks as primitives. Against this, Pauli insists that the theory is not a purely phenomenological theory, which needs to take rods and clocks as fundamental objects, and that the theory can recognize that a rod is a complicated microphysical object, whose micro-dynamics is governed by the theory of relativity. But by the same token—and this is the point of the second quote—the theory is not constructive and its value consists precisely in the fact that it shows that paradigmatically relativistic effects are derivable without making any specific assumptions about the structure of matter and the dynamics governing it.

Where does this leave the debate between Brown and Janssen? In taking Brown to argue that properly interpreted relativistic explanations of time dilation and length contraction need to invoke the details of the micro-dynamics, Janssen appear to have been misled by Brown's attempt to fit his interpretation into Einstein's principle–constructive framework. In fact both agree that the

[4] Isn't it better to avoid the issue of explanation completely in this context, as Norton has urged, since just what it means to explain only leads to "futile disputes" (Norton, 2008, p. 824)? But despite his announcement that he will "eschew explanatory issues" Norton's own final argument against Brown centrally involves an appeal to Janssen's account of common origin explanations (Norton, 2008, Section 6). I will criticize Janssen's strategy in Section 6 below.

'big principle' of Lorentz invariance is all that is needed in order to account for these phenomena. According to Brown special relativity provides, in the principle of Lorentz invariance, a universal constraint on the nature of the non-gravitational interactions (Brown, 2005, p. 147). According to Janssen, Lorentz invariance is a constraint that "transcends the individual laws" (Janssen, 2009, p. 28). So to the extent that both Brown and Janssen take the theory of relativity to postulate a universal constraint on physical theories, there is no disagreement between the two.

5. The kinematics–dynamics distinction

My discussion so far has focused on the distinction between principle and constructive theories. A second distinction which Brown and Janssen invoke to express their views is the distinction between kinematics and dynamics: Brown argues that relativity provides dynamical explanations while Janssen claims that the explanations are kinematical. Janssen argues that this constitutes a real disagreement between him and Brown. Thus, he says: "I have argued that the main objection against Lorentz's theory is that it seeks to provide dynamical explanations for a class of phenomena, namely all manifestations of Lorentz invariance, that special relativity revealed to be kinematical. That objection also applies to Brown's proposal" (Janssen, 2009, p. 27), and he emphasizes this point again later, when he introduces his case studies: "As the examples discussed in Sections 2–4 illustrate, many phenomena that would be classified as dynamical both in Lorentz's theory and in Brown's proposal are reclassified as kinematical in special relativity." But here, too, Brown's and Janssen's uses of the distinction suggest a disagreement where there actually is none.

For Janssen, a phenomenon is kinematical, "if it is just an instance of some generic feature of the world" (Janssen, 2009, p. 27) and "independent of the details of the dynamics" (31) and hence the principle of Lorentz invariance for Janssen is a *kinematical* constraint. For Brown, by contrast, Lorentz invariance is a *dynamical* constraint, yet for him a constraint is dynamical if it is a constraint on the form of any dynamics—or, in the case of the principle of Lorentz invariance "a universal constraint on the nature of the non-gravitational interactions" (Brown, 2005, p. 147). Since the principle, in virtue of providing a universal constraint, captures a default or generic feature of the world, Brown's and Janssen's claims are not incompatible. Brown insist that the default behavior of objects captured in the theory of relativity follows from a universal constraint on the dynamics, but that is compatible with the claim that the phenomena in question are instances of some generic feature of the world—that is, with the claim that the phenomena are kinematical in Janssen's sense. To the extent that the case studies Janssen discusses show that relativistic phenomena are kinematical in the sense of being independent of the details of the dynamics, they show something with which Brown, I take it, would be happy to agree.

Hence, the disagreement between Brown and Janssen concerning the classification of relativity as dynamic or kinematic seems to be one about terminology rather than about substance. Despite the passages I cited above, it seems that at least at one point Janssen seems to agree with this characterization, when he says that his "disagreement with Brown is therefore ultimately about how to draw the line between kinematics and dynamics in special relativity" (Janssen, 2009, p. 27). If the disagreement is ultimately about how to draw a conceptual distinction, then the fact that Brown and Janssen characterize the status of Lorentz invariance differently does not imply that there is a substantive disagreement between them.

Janssen (2009) also introduces a further distinction between two senses of kinematics—a broad and a narrow sense. Kinematics in the narrow sense specifies the generic spatio-temporal behavior and Janssen maintains that relativistic phenomena are kinematic in the broad sense because they are kinematic in the narrow sense. But it seems to me that even the classification of phenomena such as length contraction and the velocity dependence of mass as kinematic in the narrow sense does not on its own introduce a disagreement, for surely the principle of Lorentz invariance as universal constraint concerns the spatio-temporal behavior of objects. If the dynamical laws are Lorentz-invariant, then there are frames in which they take their canonical form and that are related by the Lorentz transformations, from which we can derive phenomena length contraction and other relativistic phenomena. Minkowski spacetime provides the natural way to encode or represent these laws, since its isometries correspond to the Lorentz transformations (see Huggett, 2009 for a summary of 'dynamic view' along these lines). Thus, the default behavior of objects expressed in the principle of Lorentz invariance centrally concerns the spatio-temporal relations among a theory's objects. Nevertheless, Janssen's notion of kinematics as specifying default spatio-temporal behavior does point to a genuine disagreement between him and Brown—not one concerning the status of relativity as being concerned with the default behavior of objects, but rather one concerning the explanatory relations between the dynamical constraint and its geometrical representations.

6. The explanatory arrow

I have argued that Brown and Janssen agree on at least one overarching lesson we can draw from Janssen's case studies—namely that the general constraint of Lorentz invariance is a common source of all instances of relativistic behavior. Janssen, however, wants to go beyond that and maintain that the structure of Minkowski spacetime provides what he calls "a common origin explanation" of universal Lorentz invariance. For Brown, by contrast, the appropriate geometric structure in which to represent the motion of physical objects is Minkowski space-time, precisely because the dynamical laws are Lorentz invariant. Here, it seem to me, we have finally reached a point of genuine disagreement.

Janssen (2009, p. 28) maintains that his case studies show not only that special relativity "imposes the kinematical constraint that all dynamical laws must be Lorentz-invariant", but that the Lorentz invariance of the laws is in turn explained by the fact that spacetime is Minkowskian. Yet it is not clear that the case studies do establish this further claim. The first case study concerns the Fresnel drag coefficient and Laue's derivation from the relativistic theorem for the addition of velocities, which does not require a detailed microscopic model. The second study traces the developments that led to the realization that measurements of the velocity dependence of the mass of the electron provide no information about the structure of the electron but only reveal a generic feature of relativistic systems. Similarly, as Janssen explains in his third case study, relativistic analyses of the Trouton–Noble experiment show that the balance of torques is a generic consequence of the definition of the four momenta and does not require any specific dynamic assumptions.

The common lesson that we can draw from these case studies appears to be that phenomena that at some point were thought to require an explanation in terms of particular hypotheses about the structure of matter were shown to be explainable independently of any details of the relevant dynamics merely by appealing to the completely general constraint of Lorentz invariance. But it is not obvious how the case studies also support the further claim that the principle of Lorentz invariance in turn has to be explained in terms of the geometry of spacetime. Take the first case study: that the Fresnel drag coefficient can be derived from the relativistic theorem for the addition of velocities shows that we need to postulate a Lorentz invariant dynamics but does not show anything about the direction of the explanatory arrow between Lorentz invariance and spacetime structure, about which Brown and Janssen disagree.

To establish the primacy of Minkowski spacetime we need an additional philosophical argument, such as Janssen's appeal to common origin explanations. Janssen's argument is this. No property of our dynamical theories except for Lorentz invariance is needed to account for phenomena such as length contraction and the velocity dependence of mass. Without further explanation, however, the fact that all our dynamical laws are Lorentz-invariant appears as a "*cosmic coincidence*" (Janssen, 2009, p. 48)—Lorentz invariance turns out to be a "property shared *accidentally*" by all dynamical laws in the absence of gravity. Yet, what appears as a coincidence on Brown's account, can in fact be explained by the statement that spacetime is Minkowskian. That is, the Minkowskian nature of spacetime provides the common origin of the fact that all dynamical laws are Lorentz-invariant. By the same token, accepting that spacetime is Minkowskian can serve as a heuristic guide for theory construction: if the Lorentz invariance of our laws were simply a gigantic coincidence, we would have no reason to expect that successful newly developed theories also ought to be Lorentz-invariant. But if we presuppose Minkowski spacetime, then we can impose Lorentz invariance as a (defeasible) constraint on any future theories as well.

We can distinguish two claims in this argument. First, Janssen claims that without further explanation the fact that all dynamical laws are Lorentz-invariant would appear to be a gigantic coincidence and, hence, this fact is in desperate need of an explanation. Second, he claims that Minkowski spacetime can in fact provide the needed explanation. I want to discuss these two claims in turn.

Brown and Pooley maintain that we can postulate "the Lorentz covariance of all the fundamental laws of physics [as] an unexplained brute fact" and they insist that "this, in and of itself, does not count against the interpretations: all explanation must stop somewhere" (Brown & Pooley, 2006, p. 80). Similarly, they would presumably argue that once we have discovered this putative 'brute fact', we could use it as a guide to future theorizing even without any further explanation. But, one might be tempted to retort, while surely every explanation must stop somewhere, the gigantic coincidence that law after law turns out to be Lorentz-invariant is precisely the kind of fact that does cry out for an explanation of the kind proposed by Janssen in terms of a common origin.[5]

Consider, however, the following analogy: two electrically charged objects exert a Coulomb force on each other. In fact, we find the same $1/r^2$-dependence for the force between charged objects in case after case. Isn't it a cosmic coincidence, one might ask, that we always observe the same kind of force between different pairs of charged objects at rest? Yet in reply to *this* question we would presumably say that this is not a coincidence at all; rather it is a matter of physical law that charged objects exert a Coulomb force on each other. We expect similar pairs of charged objects to exert an electrostatic force in accordance with Coulomb's law on each other and we use this law as a useful constraint in further modeling of the behavior of charged objects. By contrast, if I noticed that every time I play a game of dice, the first number I roll is a six, then I would regard this to be a rather puzzling coincidence and I would probably be not very likely to use this regularity as a heuristic guide to model future die rolls.

My point here is that there is at least one use of the notion of coincidence, according to which we distinguish phenomena that we can model and, hence, explain by appealing to scientific laws from those that we cannot model in this way and that, therefore, strike us as being a matter of coincidence. If we have reasons to believe that a phenomenon can be explained by appealing to physical laws, then this in *itself* is a reason enough to think of the phenomenon as not being coincidental. This distinction tracks the distinction between what we do and do not take to be projectable and to provide a heuristic constraint on future modeling.

Arguably, the 'big principle' of Lorentz invariance plays a role analogously to Coulomb's law in my example, only 'one level up', as it were—that is, as a meta-law or general constraint on the laws, rather than as a law governing the phenomena. Just as the discovery of Coulomb's law allows us to see the mutual electrostatic attraction between charged pairs of objects as not being a matter of coincidence but as a consequence of a physical law, the discovery of the principle of Lorentz invariance allows us to see the Lorentz invariance of our dynamical theories not as a coincidence but as a consequence of a more general constraint. Discovering a general meta-nomological constraint on the dynamical laws, be it Lorentz invariance or the satisfaction of energy–momentum conservation, is itself reason to think of the common feature as not being a mere coincidence. The principle of Lorentz invariance functions as a 'super law' in Eugene Wigner's sense, which "provides a structure or coherence to the laws of nature just as the laws provide a structure and coherence to the set of events" (Wigner, 1964, pp. 16 and 17). But then that principle can serve as the 'endpoint' of an explanation of a common structure of our dynamical laws, just as the laws can serve as endpoints of explanations of the common structure of the phenomena.

I want to stress that I am not arguing here that requests for explanation *must* stop once we have explained a phenomenon by appealing to a physical law (or that we can never explain meta-nomological constraints in terms of further considerations). Rather I am arguing that once we are able to explain a common feature of different phenomena in terms of the physical laws governing these phenomena, we are *not required* to offer a further explanation; similarly, once we have found a common constraint on physical theories, we are not required to offer a further explanation of this constraint. Thus, my point is merely that the realization that all our dynamical laws are Lorentz-invariant does not automatically generate a demand for a further explanation of that principle, as Janssen suggests; just as the realization that all pairs of charged objects are subjected to a Coulomb force does not automatically generate a demand for a further explanation of the Maxwell–Lorentz equations.

There is one train of thought, however, that might suggest the need for further explanation—a train of thought familiar from anti-Humean considerations. If, with the Humean, we think of laws merely as true universal generalizations, we might wonder what ensures that the generalization continues to hold. What makes it the case, or what brings it about, that future pairs of charged objects will equally be subjected to a Coulomb force, one might ask. The non-Humean will here appeal to a notion of nomological necessity that goes beyond the statement of a mere regularity and that is meant to explain why the regularities hold. It is not enough that the regularities are mere "accidental" regularities; rather, in order for them to have explanatory force, they have to be nomologically necessary. If one finds some such intuitions compelling, one might be inclined to think that merely pointing to a common property of the laws does not explain why the laws have that property (just as pointing to Coulombs law as mere regularity does not explain why charged objects behave the way they do) and the fact that all our laws share the property of Lorentz invariance must seem like a 'cosmic coincidence' unless

[5] Norton (2008, Section 6) invokes the same argument.

this can be explained in terms of some meta-law that carries a higher grade of necessity than the laws themselves do.[6] A proposal along these lines, positing a hierarchy of grades a necessity, is spelled out in Lange (2007), but I suspect that neither Janssen nor Brown would like to endorse the metaphysical intuitions driving such a project. Unless, however, we do adopt some such substantive metaphysical assumptions, the explanatory demand Janssen identifies does not genuinely arise.

Even though the principle of Lorentz invariance might not *require* a further explanation, it might turn out that such an explanation, in terms of an appeal to spacetime geometry, *can* as a matter of fact be given. Does the geometry of Minkowski spacetime provide such an explanation? I believe that the answer is no, but I have nothing substantially new to add to the arguments in Brown (2005) and Brown and Pooley (2006). One direction such an explanation might take is that it appeals to Minkowski spacetime as an independently existing entity and argues that the metric structure of this spacetime constrains the laws to have a certain form. Now, as we have already seen above Brown rejects this kind of explanation, calling the question how rods and clocks might be able to know what spacetime they are immersed in the ''mystery of mysteries'' (Brown, 2005, p. 143).

This is not, however, the kind of explanation Janssen proposes, who instead endorses (at least for the purposes of the argument) Brown's spacetime relationalism and the view that spacetime only serves to codify key aspects of the behavior of particles. But, as Brown and Pooley have argued (Brown & Pooley, 2006, p. 81), it is puzzling how Minkowski spacetime can merely serve to ''encode'' the default spatio-temporal behavior of physical systems and yet at the same time also provide an *explanation* of the Lorentz-invariance of the laws. Indeed, it would seem that whatever property is encoded is explanatorily prior to the structure in which it is encoded. Janssen suggests that the statement that spacetime is Minkowskian goes beyond the statement that all laws are Lorentz-invariant by explaining all 'relativistic phenomena' ''in one fell swoop''. But the principle of Lorentz invariance, understood as *universal* constraint on any dynamical theory, likewise explains all these phenomena in one fell swoop, tracing them to a common origin—'the big principle'. Similarly, the principle of Lorentz invariance ''yields a negative and a positive heuristic'' (Janssen, 2009, p. 49). It tells scientists that certain classes of phenomena can be explained by appealing to this principle alone, without knowledge of the details of the relevant dynamics, and it provides ''useful constraints on further theorizing about elements in the class of dynamical phenomena''. Thus it is not clear what explanatory resources the claim that spacetime is Minkowskian, made within a relationalist framework, can add to the principle of Lorentz invariance.

I said that both Brown and Janssen might be hesitant to endorse a rich, non-Humean metaphysics. Yet 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 & Pooley, 2006, p. 77).

[6] But we need to be careful about how to put the point: Once we take the laws to be nomologically necessary, no feature of the laws can be truly 'accidental' or a 'coincidence'. Instead we would have to distinguish those features that are merely nomologically necessary from those that follow from some higher-level meta-nomological constraint.

References

Balashov, Yuri, & Janssen, Michel (2003). Presentism and relativity. *British Journal for the Philosophy of Science*, *54*, 327–346.

Brown, Harvey, & Pooley, Oliver (2006). Minkowski space–time: A glorious non-entity. In D. Dieks (Ed.), *The ontology of spacetime* (pp. 67–92). Amsterdam: Elsevier.

Brown, Harvey R. (2005). *Physical relativity: Space–time structure from a dynamical perspective*. Oxford: Oxford University Press.

Einstein, Albert (1919). Time, space, and gravitation. *The London Times*, November 28. Page references to reprint (under the title: What is the theory of relativity?) In A. Einstein, *Ideas and opinions* (pp. 227–232). New York: Crown Publishers, 1954.

Frisch, Mathias (2005). Mechanisms, principles, and Lorentz's cautious realism. *Studies in History and Philosophy of Modern Physics*, *36*(4 December), 659–679.

Huggett, Nick (2009). Physical relativity: space-time structure from a dynamical perspective by Harvey Brown; Understanding space-time: The philosophical developments of physics from Newton to Einstein by Robert DiSalle. *Philosophy of Science*, *76*(3), 404–422.

Janssen, Michel (2009). Drawing the line between kinematics and dynamics in special relativity. *Studies in History and Philosophy of Modern Physics*, *40*, 26–52.

Lange, Marc (2007). Laws and meta-laws of nature: Conservation laws and symmetries. *Studies in History and Philosophy of Modern Physics*, *38*, 457–481.

Lorentz, H. A. (1900). *Elektromagnetische Theorien physikalischer Erscheinungen*. *Collected papers*, Vol. VIII. The Hague: Martin Nijhoff.

Norton, John (2008). Why constructive relativity fails. *British Journal for the Philosophy of Science*, *59*, 821–834.

Wigner, Eugene (1964). Symmetries and conservation laws. In *Symmetries and reflections*. Greenwood Press.