Western Graduate&PostdoctoralStudies

Western University
Scholarship@Western

Electronic Thesis and Dissertation Repository

September 2013

Some disputed aspects of inertia, with particular reference to the equivalence principle

Ryan S. Samaroo
The University of Western Ontario

Supervisor
Professor Robert DiSalle
The University of Western Ontario

Graduate Program in Philosophy

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

© Ryan S. Samaroo 2013

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Philosophy of Science Commons

Recommended Citation
Samaroo, Ryan S., "Some disputed aspects of inertia, with particular reference to the equivalence principle" (2013). *Electronic Thesis and Dissertation Repository*. 1619.
https://ir.lib.uwo.ca/etd/1619

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca.

SOME DISPUTED ASPECTS OF INERTIA,
WITH PARTICULAR REFERENCE TO THE EQUIVALENCE PRINCIPLE

(Thesis format: Integrated article)

Ryan Samaroo

Graduate Programme in Philosophy

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

The School of Graduate Studies
University of Western Ontario

© Ryan Samaroo 2013

Abstract

This thesis is a contribution to the foundations of space-time theories. It examines the proper understanding of the Newtonian and 1905 inertial frame concepts and the critical analysis of these concepts that was motivated by the equivalence principle. This is the hypothesis that it is impossible to distinguish locally between a homogeneous gravitational field and a uniformly accelerated frame.

The three essays that comprise this thesis address, in one way or another, the criteria through which the inertial frame concepts are articulated. They address the place of these concepts in the conceptual framework of physics and their significance for our understanding of space and time.

In Chapter 2, I examine two claims that arise in Brown's (2005) account of inertia. Brown claims there is something objectionable about the way in which the motions of free particles in Newtonian theory and special relativity are coordinated. Brown also claims that since a geodesic principle can be derived in Einsteinian gravitation the objectionable feature is explained away. I argue that there is nothing objectionable about inertia and that, while the theorems that motivate Brown's claim can be said to figure in a deductive-nomological explanation, their main contribution lies in their explication rather than their explanation of inertial motion.

In Chapter 3, I examine Friedman's recent approach to the analysis of physical theories (2001; 2010a; 2010b; 2011). Friedman argues that the identification of certain principles as 'constitutive' is essential to the correct methodological analysis of physics. I explicate Friedman's characterisation of a constitutive principle and his account of the constitutive principles that Newtonian and Einsteinian gravitation presuppose for their formulation. I argue that something very close to Friedman's view is defensible.

In Chapter 4, I examine the so-called background-independence that Einsteinian gravitation is said to exemplify. This concept has figured in the work of Rovelli (2001; 2004), Smolin (2006), Giulini (2007), and Belot (2011), among others. I propose three ways of fixing the extension of background-independence, and I argue that there is something chimaerical about the concept. I argue, however, that there *is* a proposal that clarifies the feature of Einsteinian gravitation that motivates the concept.

Keywords: inertial frames; Newtonian gravitation; equivalence principle; geodesic principle; general relativity; background-independence; structure of space-time theories; theory change; methodology

Acknowledgements

I wish to thank Professors Robert DiSalle, Bill Demopoulos, and Wayne Myrvold for their generous help and constant support, for their wisdom and their friendship. This thesis grew out of conversations with them.

I would like to thank Professor Barry Hoffmaster for his good advice and support.

I am grateful to the Warden and Fellows of Merton College, Oxford, for their warmth and welcome in the 2009-2010 academic year; and my thanks to Dr David Wallace and the Philosophy of Physics Group.

I gratefully acknowledge support from the Social Sciences and Humanities Research Council of Canada.

None of this would have been possible without my family, whose support of my studies, and everything else, has given me the freedom to pursue my interests to the fullest. Last, I want to thank my wife Ariella Binik for all her encouragement and for making many helpful suggestions.

Table of Contents

Chapter 1

Introduction

This thesis consists of three essays. Each essay examines a view that has been defended by a recent interpreter of relativity. The essays share a concern with the proper understanding of the Newtonian and 1905 inertial frame concepts and the critical analysis of these concepts that was motivated by an insight Einstein had in 1907. That insight has been called ‘the equivalence principle,’ and it is the hypothesis that it is impossible to distinguish locally between a homogeneous gravitational field and a uniformly accelerated frame.[1]

The equivalence principle motivated what I call ‘the 1907 inertial frame concept’ and, with it, a framework of investigation in which Einstein could construct his gravitation theory. The concept applies to the state of motion of a reference frame that falls freely. That is to say, the frame is subject only to gravitation and has no other component of acceleration; for example, it has no component of rotation. In a sufficiently local and homogeneous region of space-time, such a frame is indistinguishable from an inertial frame in special relativity.

The inertial frame concepts at issue—Newtonian, 1905, 1907—arise in Newton’s and Einstein’s theories. These theories provide a framework of constraints for theories of special systems such as perfect fluids or gravitation. These ‘framework-theories’ are founded on general principles of nature that physical processes satisfy and that motivate

[1]In my discussion of the principle, I have drawn heavily on the treatments of Anderson (1967) and Ehlers (1973); and there are still other ways of characterising the equivalence principle. See, e.g., the taxonomy of Will (1993). Nonetheless, when I write ‘the equivalence principle,’ I am referring to Einstein’s insight of 1907 that it is impossible to distinguish locally between a homogeneous gravitational field and a uniformly accelerated frame. This principle has been called ‘Einstein’s equivalence principle’ or ‘the 1907 equivalence principle,’ and it is an interpretive extrapolation from what Anderson, Ehlers, and Will would call ‘the principle of the universality of free fall’ or ‘the weak equivalence principle.’ While the principles of Anderson, Ehlers, and Will are formulated in the context of Einstein’s completed theory of gravitation, the 1907 equivalence principle is situated in the context of theory construction. Isolating it is important for understanding the identification of inertial frames and freely falling ones that was crucial to Einstein’s construction of his gravitation theory.

mathematically formulated *criteria*.[2] These criteria enable us to articulate theoretical concepts such as force, mass, acceleration, rotation, and simultaneity; and through the articulation of these concepts, they bear on the decomposition of motions into their inertial and non-inertial components. In this way, the inertial frame concepts are tied to the development of the general principles. These essays examine the general principles and the criteria these suggest—their provision, their methodological function, their place in an account of the structure of theories. The essays share, moreover, a preoccupation with the geometric objects whose application the criteria control, in particular, with the transition from those geometric objects peculiar to Newtonian theory and special relativity to those peculiar to Einsteinian gravitation.

Each essay takes as its starting point the Newtonian inertial frame concept, and it is helpful to review some of the early history of that concept. This will bring to light a few turning-points and features that remain with us, and will situate the essays in a broader tradition.

1.1. Groundwork for a perturbative analysis of motion

If one wanted to go very far back, one could begin with the concept of natural motion. Its history is well documented, from Aristotle and the Epicureans, the early and later impetus theorists, to the Early Moderns. But those episodes that I will emphasise concern the recognition of a concept of uniform motion that began in the seventeenth century. This concept provided the groundwork for a perturbative analysis of motion, one that could serve as the basis for a theory by which interactions in a physical system could be both observed and measured. Three figures should be singled out: Galileo, Huyghens, and Newton.

Galileo was the first to formulate a number of interrelated concepts that are essential to the analysis of a system of bodies in motion. They arise in his *Dialogue*

[2]This sketch of a distinction between theories that provide a general framework for physics and specific theories constructed within such a framework is drawn from Einstein's (1919 [2002]) distinction between principle and constructive theories, but informed by the interpretations of Flores (1999) and DiSalle (2006; 2012), the latter informed by Demopoulos (1974) and Bub (2005).

Concerning the Two Chief World Systems (1632 [1967]). While the *Dialogue* covered a great deal of contemporary science, these concepts are articulated in the passages directed against the conception of motion peculiar to the Aristotelian and Ptolemaic systems. Aristotle defended a concept of 'natural motion,' which he understood with reference to 'natural place.' In the terrestrial realm, the four earthly elements—earth, air, fire, water—move in straight lines towards their natural places unless they are hindered. For example, since smoke is mostly air it moves naturally towards the sky, but not as high as fire; and heavy bodies move naturally towards the centre of the universe. Aristotle held that objects deviate from their natural motions only when forced, and then only while the force is being applied. In this way, he drew a distinction between natural and unnatural or forced motions. In the celestial realm, the heavenly bodies move through the incorruptible aether, and so move perfectly and eternally in the celestial spheres around the stationary Earth. (*Physics* IV, 1; *On the heavens* I, 2)

In the *Dialogue*, Simplicio is an Aristotelian proponent of the geocentric hypothesis; Sagredo is an intelligent and initially neutral layman; and Galileo's spokesman is Salviati. Their debate is over the geocentric and heliocentric hypotheses. Simplicio gives an argument, now known as 'the tower argument,' to undermine Copernicus' heliocentric hypothesis that the Earth makes diurnal rotations about its axis and annual revolutions around the stationary Sun. This argument is directed at the claim that the Earth makes diurnal rotations:

> For, if [the Earth] made the diurnal rotation, a tower from whose top a rock was let fall, being carried by the whirling of the earth, would travel many hundreds of yards to the east in the time the rock would consume in its fall, and the rock ought to strike the earth that distance away from the base of the tower. (Galileo, 1632 [1967], p. 126)

Salviati points out that the argument begs the question of the rock's vertical motion. He challenges the assumption that the vertical motion of the rock, from the top of the tower to the ground, cannot be composed with another motion, the circular motion described by the tower. He argues that the rock participates in the motion of the Earth and stays with the tower. That is to say, the motion of a falling rock is 'not straight at all, but mixed straight-and-circular.' (Galileo, 1632 [1967], p. 248) The principle that underlies the fact that the rock stays with the tower has been called the *principle of the composition of motions*.

The composition of motions is a significant insight. It motivates a critical analysis of the Aristotelian idea that terrestrial bodies move in straight lines towards their natural places unless they are hindered. The new concept of natural motion that is revealed in this analysis can be summarised in a *principle of uniform circular motion*: A body perseveres in its state of being at rest or of moving in uniform circular motion unless it is disturbed. We find in Galileo, therefore, the first clear recognition of a state of uniform motion relative to which other motions can be referred.

The Galilean composition of motions—in particular, the idea that each component of a motion is fully realised—represents a significant break with the Aristotelian intuition that the motions of terrestrial bodies must oppose one another. Simplicio opposes the idea that the observation statement 'the stone is falling straight down' should be reinterpreted in light of the new concept of composition that Salviati is defending. Salviati tries to convince him of the composition of motions by showing him that it is already implicitly in use. He proceeds to explicate the concept by taking a familiar example of composed motion that goes unnoticed: Sailors below deck in a large, moving ship notice that smoke from burnt incense does not move backwards, behind the ship, but rises in a column in the same way that it would at the dock; the motion of the smoke is composed with the motion of the ship. Likewise, jumping in place aboard the moving ship does not result in the ship passing underneath; the motion of the person jumping is composed with the motion of the ship. The enclosed cabin of the ship can be generalised to any space in sufficiently uniform motion for accelerations to be undetected, one furthermore that permits the geometrical description of the motions of the bodies among themselves.[3] Such a space can be called a *reference frame*. With this concept, the lesson of the ship can be summarised in the *Galilean relativity principle*: Given a system of bodies in a

---

[3]From a retrospective point of view, we might say that such a space is sufficiently local, homogeneous, and isotropic, one furthermore in which an accelerometer would detect no acceleration.

reference frame, their motions relative to one another are the same whether the frame is at rest or moving uniformly.[4] The notions of a reference frame and a relativity principle, when they are taken together, determine an *equivalence class of reference frames* in uniform motion.

Having convinced Simplicio that the principle of the composition of motions is already implicitly in use, Salviati proceeds to convince him that, in virtue of the principle, a rock dropped from the mast of a moving ship would not be left behind but would fall at the base of the mast. He then argues that the principle applies just as well to the tower as to the ship. The motion of the rock falling from the tower and the motion of the rock falling from the mast are both instances of composed motion:

> transfer this argument [the ship argument] to the whirling of the earth and to the rock placed on top of the tower, whose motion you cannot discern because, in common with the rock, you possess from the earth that motion which is required for following the tower; you do not need to move your eyes. Next, if you add to the rock a downward motion which is peculiar to it and not shared by you, and which is mixed with this circular motion, the circular portion of the motion which is common to the stone and the eye continues to be imperceptible. The straight motion alone is sensible, for to follow that you must move your eyes downwards. (Galileo, 1632 [1967], p. 250)

Salviati's essential point in all of these arguments is this: The fact that the rock falls at the base of the tower does not establish anything either way; it does not undermine the argument that the rock's motion is 'mixed straight-and-circular.' What emerges from the dialogue is that the concepts of composed motion and uniform circular motion, far from being radical, are in fact already implicitly in use.[5]

Setting aside Galileo's goal of defending the Copernican hypothesis, what is of greatest interest to us is his recognition of a number of interrelated concepts. We find in

---

[4]Galileo's clearest statement of his principle can be found in the 'Second Day' of the *Dialogue* (1632 [1967], p. 187). This statement of Galileo's relativity principle is close to Newton's Corollary V to the laws of motion (Newton, 1726 [1999], p. 423). It differs in that it makes no mention of rectilinearity. Note also that, for Galileo, the principle is an independent principle; for Newton, it is a consequence of the laws.

[5]To emphasise the Socratic aspect of the *Dialogue*, one could say that these concepts already reside within us—within Simplicio and Sagredo—and we are brought by anamnesis—by Salviati's midwifery—to recollect them.

'the tower argument' the *principle of the composition of motions* and the *principle of uniform circular motion*. We find in 'the ship argument' the first clear notions of a *reference frame*, a *relativity principle*, and an *equivalence class of reference frames*. What Galileo recognised, in sum, is the concept of a reference frame in a privileged state of motion, one relative to which we can give a satisfactory description of the motions among the bodies.

While the debate over the geocentric and heliocentric hypotheses continued, the mechanical philosophers focused on the more specific, though related, problem of giving a proper analysis of collisions. Huyghens understood certain aspects of the relativity of motion better perhaps than any of his contemporaries, and certainly better than Descartes and Leibniz.[6] He saw clearly that place, rest, motion, velocity, and acceleration are all relative to some arrangement of bodies taken as a reference—to a table, e.g., on whose surface a number of balls are rolling. He recognised that velocity-difference and *not* absolute velocity is an objectively measurable quantity. This is especially clear in his analysis of rotation. Rotation might appear to stand apart from other relative motions since in involves no change of position. But Huyghens recognised that rotation is not absolute but a species of relative motion; it is peculiar because it is a property of a system of particles—the smallest rotating body must have at least two points with different velocities. Thus, rotation amounts to the relative velocity of the two points.

On the surface, Huyghens shared with Descartes and Leibniz the view that all motions are relative. But where Descartes and Leibniz defended confused and sweeping accounts of relative motion—accounts on which *any* state of motion is essentially as good as any other for saying which bodies are in motion and which at rest—Huyghens recognised that determining places and velocities, accelerations and rotations, implicitly depends on a *privileged* state of uniform rectilinear motion relative to which they can be referred. Huyghens saw clearly that such a state of motion is needed for a satisfactory

---

[6]My discussion of Huyghens follows closely that of Stein (1967; 1977), who has translated previously unpublished fragments of Huyghens, and DiSalle (2006).

description of motion and for analysing the exchange of a 'quantity of motion' or momentum in collisions.

In *Principia* (1726 [1999]), Newton took the concept of a reference frame in uniform rectilinear motion and generalised it to comprehend any frame relative to which accelerations can be understood as the result of the action of some force. Newton's laws of motion express criteria for applying the concept of force, and they allowed Newton to distinguish uniform rectilinear motion or *inertial motion* from accelerated or forced motion. The first law of motion or principle of inertia defines this relation between motion and force:

> *Every body perseveres in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by forces impressed.* (Newton, 1726 [1999], p. 416)

With the principle, we have moved beyond a merely kinematical notion of a reference frame—a frame sufficient for a geometrical description of motion without any appeal to causes—to a new one that is inseparable from Newton's account of force.

But Newton's account of inertial motion does not end with the statement of the first law: The third law of motion is essential to that account—a point I will examine in detail in Chapter 1. When they are taken together, Newton's laws identify a privileged class of dynamically equivalent reference frames satisfying Galileo's relativity principle. In any such frame, forces and masses, accelerations and rotations, have the same measured values. However, much as these frames are empirically indistinguishable, for Newton, they were not theoretically equivalent: Newton thought of them as moving with various velocities relative to what he called 'absolute space,' even though those velocities cannot be known.

Though 'absolute space' was a technical term in Newton's theory, and so departs from the established meaning, many of his contemporaries understood it to invoke a metaphysical thesis and criticised its introduction on philosophical grounds. It was not until the nineteenth century that Newton's theory was given its proper form by the insight into its complete independence from the notion of absolute space in the work of Neumann (1870), Thomson (1884), Lange (1885), and others.[7] Following this work, the content of the laws of motion can be summarised: Given a system of particles in motion, there exists a reference frame and a time-scale relative to which every acceleration is proportional to and in the direction of the force applied, and where every such force belongs to an action-reaction pair.[8]

The laws of motion so understood give rise to the inertial frame concept that Einstein subjected to a critical analysis, in two stages. In the 1905 paper, Einstein argued from the nineteenth-century concept to a new one motivated by an electrodynamical principle. The new concept to which the nineteenth-century one gave way is this: An inertial frame is not merely one in uniform rectilinear motion but one in which light travels equal distances in equal times in arbitrary directions. It was *this* concept of motion that Einstein subjected to a critical analysis in 1907 through his insight—what is often called 'the equivalence principle'—that it is impossible to distinguish locally between a homogeneous gravitational field and a uniformly accelerated frame. The equivalence principle motivates the critical analysis of the 1905 concept precisely because classical inertial frames cannot be distinguished locally from freely falling ones.[9]

My project examines three disputed ideas that are bound up with this last episode. It examines a recent account of the *classical inertial frame concept*; it examines an account of the *equivalence principle*; and it examines an alleged methodological lesson of the revision of the *background-structure* peculiar to Newtonian theory and special relativity.

1.2. Brief outline

In Chapter 2, I examine two claims that arise in Harvey Brown's account of inertial motion in *Physical Relativity* (2005). Brown claims there is something

---

[7]Other notable contributors include Mach (1883 [1919]), Muirhead (1887), and MacGregor (1893).
[8]This formulation is close to those of Thomson (1884, p. 387) and Muirhead (1887, pp. 479-480).
[9]It is noteworthy that this critical analysis can be applied separately to the Newtonian and 1905 inertial frame concepts, since, independently of electrodynamics, Newtonian forces look the same in freely falling frames (cf. Corollary VI to the Laws of Motion).

objectionable about the way in which the motions of free particles in Newtonian theory and special relativity are coordinated. This claim implies that inertia requires an explanation since the coordination is postulated not explained. Brown also claims that since a geodesic principle can be derived as a theorem in Einsteinian gravitation the objectionable feature of Newtonian theory and special relativity is explained away. I take issue with both claims. I argue that there is nothing objectionable about inertia and that, while the theorems that motivate Brown's claim can be said to figure in a deductive-nomological explanation, their main contribution lies in their explication rather than their explanation of inertial motion.

In Chapter 3, I examine Friedman's recent approach to the analysis of physical theories (2001; 2010a; 2010b; 2011). Friedman argues against Quine that the identification of certain principles as 'constitutive' is essential to the correct methodological analysis of physics. I explicate Friedman's characterisation of a constitutive principle and his account of the constitutive principles that Newtonian and Einsteinian gravitation presuppose for their formulation. I argue that something very close to Friedman's view is defensible.

In Chapter 4, I examine the so-called background-independence that Einsteinian gravitation is said to exemplify. A number of physicists and philosophers of physics, notably Rovelli (2001; 2004) and Smolin (2006), have taken this background-independence to be an insight about nature that ought to be preserved in a future theory, and they take the pursuit of background-independent approaches to quantum gravity to be a good heuristic. I ask: What is this sought-after background-independence? What is the concept at stake here? I propose three ways of fixing the extension of background-independence, and I argue that there is something chimaerical about the concept. I argue, however, that there *is* a proposal for background-independence that clarifies the feature of Einsteinian gravitation that is the basis for nearly all proposals for background-independence.

In Chapter 5, I conclude with a brief discussion of two notions of apriority that run through the essays. I outline a way of extending this project in future work.

References

Anderson, J. L. (1967). *Principles of Relativity Physics*. New York: Academic Press.

Aristotle. (1984). *The Complete Works of Aristotle* (J. Barnes, Trans.). Princeton: Princeton University Press.

Brown, H. (2005). *Physical Relativity: Space-time Structure from a Dynamical Perspective.* Oxford: Oxford University Press.

Bub, J. (2005). Quantum mechanics is about quantum information. *Foundations of Physics, 34*, 541-560.

Demopoulos, W. (1974). What is the logical interpretation of quantum mechanics? In R. Cohen et al. (Eds.), *PSA 1974: Proceedings of the biennial meeting of the philosophy of science association* (pp. 721-728). Dordrecht: Reidel.

DiSalle, R. (2012). Analysis and Interpretation in the Philosophy of Modern Physics. In M. Frappier, D. Brown, & R. DiSalle (Eds.), *Analysis and Interpretation in the Exact Sciences: Essays in Honour of William Demopoulos* (pp. 1-18). Dordrecht: Springer Netherlands.

DiSalle, R. (2009). Space and Time: Inertial Frames. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2013). URL = <http://plato.stanford.edu/archives/win2009/entries/spacetime-iframes/>.

DiSalle, R. (2006). *Understanding Space-Time.* Cambridge: Cambridge University Press.

Ehlers, J. (1973). Survey of General Relativity Theory. In W. Israel (Ed.), *Relativity, Astrophysics, and Cosmology* (pp. 1-125). Dordrecht: D. Riedel.

Einstein, A. (2002). What is the theory of relativity? In M. Janssen & al. (Eds.), *The Collected Papers of Albert Einstein, vol. 7* (pp. 206-215). Princeton: Princeton University Press. (Original work published in 1919)

Flores, F. (1999). Einstein's theory of theories and types of theoretical explanation. *International Studies in the Philosophy of Science, 13*, 123-134.

Friedman, M. (2011). Einstein and the a priori. Unpublished manuscript.

Friedman, M. (2010a). Synthetic History Reconsidered. In M. Domski & M. Dickson (Eds.), *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science* (pp. 571-813). Chicago: Open Court.

Friedman, M. (2010b). A post-Kuhnian approach to the history and philosophy of science. *The Monist, 93*, 497-517.

Friedman, M. (2001). *Dynamics of Reason: The 1999 Kant Lectures of Stanford University.* Stanford: CSLI Publications.

Galileo. (1967). *Dialogue Concerning the Two Chief World Systems, 2nd Edition* (S. Drake, Trans.). Berkeley: University of California Press. (Original work published in 1632)

Lange, L. (1885). Ueber das Beharrungsgesetz. *Berichte der Koniglichen Sachsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-physische Classe, 37*, 333-351.

Mach, E. (1919). *The Science of Mechanics: A Critical and Historical Account of its Development* (T. McCormack, Trans.). Chicago: Open Court. (Original work published in 1883)

MacGregor, J. (1893). On the Hypotheses of Dynamics. *Philosophical Magazine, 36*, 233-264.

Muirhead, R. (1887). The Laws of Motion. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 25*, 473-489.

Neumann, C. (1870). *Ueber die Principen der Galilei-Newton'schen Theorie*. Leipzig: B. G. Teubner.

Newton, I. (1999). *The Principia: Mathematical Principles of Natural Philosophy* (I. B. Cohen & A. Whitman, Trans.). Berkeley: University of California Press. (Original work published in 1726)

Rovelli, C. (2004). *Quantum Gravity*. Cambridge: Cambridge University Press.

Rovelli, C. (2001). Quantum Spacetime: What Do We Know? In C. Callender & N. Huggett (Eds.), *Physics Meets Philosophy at the Planck Scale*. Cambridge: Cambridge University Press.

Smolin, L. (2006). The Case for Background Independence. In D. Rickles, S. French, & J. Saatsi (Eds.), *The Structural Foundations of Quantum Gravity*. Oxford: Oxford University Press.

Stein, H. (1977). Some Philosophical Prehistory of General Relativity. In J. Earman, C. Glymour, & J. Stachel (Eds.), *Foundations of Space-Time Theories. Minnesota Studies in the Philosophy of Science, vol. VIII* (pp. 3-49). Minnesota: University of Minnesota Press.

Stein, H. (1967). Newtonian space-time. *Texas Quarterly, 10*, 174-200.

Thomson, J. (1884). On the law of inertia; the principle of chronometry; and the principle of absolute clinural rest, and of absolute motion. *Proceedings of the Royal Society of Edinburgh, 12*, 568-578.

Will, C. (1993). *Theory and Experiment in Gravitational Physics*. Cambridge: Cambridge University Press.

Chapter 2

There is no conspiracy of inertia

1. Introduction

Conceptual analysis, at least in the analytic tradition since Frege, is the practice of identifying central features of a concept by revealing the assumptions on which use of the concept depends.[10] This approach to conceptual analysis has also been a part of the foundations of physics, at least since Newton. Conceptual analysis in physics is responsible to the body of theory and practice in which the concept is situated and in which it is interconnected with other concepts both physical and mathematical. The identification and explication of these connections, therefore, is a main objective of an analysis.

There is, however, an older tradition in which conceptual analysis does not proceed in this way; concepts are explicated through metaphysical and methodological enquiry. The intuitions that drive this kind of enquiry are held to bear decisively on the analysis of physical theories and the concepts they comprise. This tradition, at least so far as the theory of space and time is concerned, is exemplified in certain arguments offered by Leibniz, Huyghens, Berkeley, Mach, and Einstein.[11] This is not to say that these thinkers did not also pursue conceptual analysis in the sense characteristic of the analytic tradition, but their most famous and influential criticisms of Newton's theory reflect underlying conceptions of substance, action, and causality that are alleged to be known independently of physics. These conceptions, furthermore, are bound up with views about how knowledge of the structure of the world is gained, about the nature of scientific explanation, and about what an empirical theory may legitimately postulate. Brown's *Physical Relativity* (2005) belongs to this tradition, and those claims that I will examine are motivated by a number of intuitions about inertial motion that have their source in Einstein. Like Einstein, Brown holds that there is something objectionable about inertial

---

[10]This way of expressing the basic idea of conceptual analysis is due to Demopoulos (2000, p. 220).

[11]This is not to say that Newton's views on space, time, and motion are free of philosophical intuitions. But Newton stands out among these thinkers for offering principles that constitute these concepts independently of any philosophical intuitions one may have about them. In this way, he ensures that the concepts are insulated from such intuitions, even his own.

motion in Newtonian theory and special relativity. He calls the objectionable feature 'the conspiracy of inertia,' and he claims that the conspiracy is explained away by Einstein's theory of gravitation.[12]

In this essay, I will examine Brown's account of inertial motion in Newtonian theory and special relativity. I will argue against the allegation of a conspiracy, and I will argue that, while there is a sense in which Einsteinian gravitation explains inertial motion, the main contribution of the theorems that motivate Brown's claim lies in their explication rather than their explanation of inertial motion.

2. The alleged conspiracy

There is a view that can be found in Einstein (e.g., 1922 [1950]; 1924 [1991]) and Nerlich (1976) according to which space-time structure explains the motion of free particles. Free particles and light rays move along the 'ruts' and 'grooves' of the affine geodesics of space-time, much as trains run along tracks. As Nerlich (1976, p. 264) has put it, it is 'because space-time has a certain shape that worldlines lie as they do.' On this view, there is a causal inference from space-time structure to the phenomena of motion. This view is sometimes called 'the space-time explanation' or 'the causal-explanatory view.'

Brown's account of space-time structure is set against this view. Brown argues that space-time structure is determined by the equivalence-class structure of some particular dynamical theory; for example, the equivalence-class structures determined by Newtonian mechanics or Maxwell's theory. Einsteinian gravitation differs from these theories because it has no symmetry group. More precisely, it has no non-trivial symmetries; its symmetries are expressed by the invariance of the geometric objects on the manifold under arbitrary diffeomorphisms. Nonetheless, elements of space-time structure are coordinated to the world of experience by an empirical criterion, namely, the

[12]These claims are separable from the principal claim of *Physical Relativity*, namely, that length contraction and clock retardation are in need of a dynamical explanation, and that such an explanation must come from a 'constructive theory,' that is, a theory of the forces of cohesion that maintain a body's configuration. This view has been criticised by Norton (2008), Hagar (2008), Janssen (2009), and DiSalle (2012).

strong equivalence principle. With the equivalence principle, approximately geodesic trajectories can be coordinated with free particles. In all of these theories, therefore, space-time structure is a codification of certain important features of physical processes. Or, to put it the other and perhaps more familiar way, certain important features of physical processes can be represented by certain elements of space-time structure.

While Brown's view is significant for its rejection of the space-time explanation account, it has another aspect that is just as problematic as the latter view.[13] For Brown, Newtonian theory and special relativity commit us to accepting something questionable, which he illustrates with the metaphor of a conspiracy among the free particles of the universe. The metaphor can be found in a number of passages, of which the following four are representative:

> Inertia, before Einstein's general theory of relativity, was a miracle. I ... mean the ... postulate that force-free (henceforth *free*) bodies conspire to move in straight lines at uniform speeds while being unable, by *fiat*, to communicate with each other. (Brown, 2005, pp. 14-15)

> A kind of highly non-trivial pre-established harmony is being postulated, and it takes the form of the claim that there exists a coordinate system x^μ and parameters τ such that [$d^2x^\mu/d\tau^2 = 0$] holds for each and every free particle in the universe. (Brown, 2005, p. 17)

> ... there is a prima facie mystery as to why objects with no antennae should move in an orchestrated fashion. That is precisely the pre-established harmony, or miracle, that was highlighted above. (Brown, 2005, p. 24)

> ... force-free particles have no antennae ... they are unaware of the existence of other particles. That is the *prima facie* mystery of inertia in pre-GR theories: how do all the free particles of the world know how to behave in a mutually coordinated way such that their motion appears extremely simple from the point of view of a family of privileged frames? (Brown, 2005, p. 142)

I propose the following as a synthesis of these and other passages that exemplify what I call

> *Brown's alleged conspiracy:* As a matter of definition, the free particles of the universe are non-interacting, and thus cannot detect other objects or even

[13] See also DiSalle (1995; 2006a) for a critique of the space-time explanation account.

> determine whether there are any.[14] Yet, they seem to move in a mutually coordinated way. How do they know to move in the way that they do? Newtonian theory and special relativity commit us to thinking that there is a conspiracy among them. These theories assert that there exists a coordinate system x^μ and parameters τ associated with each particle such that the equation $\mathrm{d}^2x^\mu/\mathrm{d}\tau^2 = 0$ holds.

To put the idea another way, one could say that the free particles of the universe agree not to accelerate and to follow geodesics of the space-time. Also, particles that are themselves composites must satisfy the conservation of momentum; that is, the forces among the constituent particles must be balanced, failing which the particle, by its internal forces, will accelerate without limit. Therefore, one could say that free particles must also conspire to maintain a state of equilibrium.[15] I will address this in detail in §5.

To take the metaphor of a pre-established harmony rather than a conspiracy, one could say that free particles are predetermined to move such that certain geometrical relations among them hold; specifically, such that the function describing the distances between them is of a certain fixed form. One could say that every free particle has as part of its complete notion that its motion will be such that the distance function holds. That is, its correspondence with other particles belongs to its complete notion and to theirs.

---

[14]Note that we are considering here only the framework of the laws of motion. In Newtonian gravitation, there is an interaction among all of the particles of the universe; every particle attracts every other with a force that is proportional to the product of their masses and inversely proportional to the square of the distance between them.

[15]Though the alleged conspiracy may be understood completely in these terms, it is worth noting that, for Brown, there must be more than *four* free particles for there to be a conspiracy. This idea is motivated by a Lange-style path-construction proposed by Pfister (2004). Pfister defines the *rectilinear aspect* of an inertial system in terms of a path-structure in projective geometry. Let three free particles emanate from an event such that they follow straight paths p, p', p'' that are non-collinear. Choose arbitrary points A, B on p, A', B' on p', and so on. Then, e_1 is the point at which the paths AA' and BB' cross, e_2 that at which AA'' and BB'' cross, and e_3 that at which $A'A''$ and $B'B''$ cross. The path $e_1e_2e_3$ is part of the structure so constructed. These four paths define the rectilinear aspect of an inertial system. But, relative to such a construction, all other free particles will move on straight lines.

While four particles are necessary to define the rectilinear aspect of an inertial system, this projective structure alone is not sufficient; it cannot satisfy the requirement that a body's motion be *uniform* with respect to time as well as rectilinear. The requirement that motion be uniform with respect to time means that a body traverses equal distances in equal times. To give an account of equal distances in equal times we must define affine and metrical structures. Only in this way do we obtain the necessary notions of parallelism and distance. While Brown's discussion of Pfister's construction goes part of the way towards explicating the mathematical requirements for inertia, it distracts from what really underlies the alleged conspiracy, namely, a preoccupation with absolute or global space-time structures.

It is worth noting the awkwardness of the conspiracy metaphor. Free particles seem to conspire in spite of the fact that they *cannot* communicate. This, presumably, is what is 'miraculous.' But, pressing on with the conspiracy metaphor, one could still ask why the particles are prohibited from conspiring to move according to some law; for example, a law relating their motion to the distribution of mass-energy.[16] Perhaps the metaphor of a pre-established harmony is more apt.

There may be better metaphors or better ways of fleshing out the existing ones with the relevant physics. But, in what follows, I am less concerned with the metaphors themselves than with the idea that any such metaphors—any *negative* metaphors—are appropriate at all. I will argue that the allegation of a conspiracy is driven by a number of metaphysical and methodological intuitions that obscure rather than clarify inertial motion in Newtonian theory and special relativity.

3. The conspirators unmasked

Brown's view belongs to a tradition according to which Einsteinian gravitation is superior to its predecessors not only because it is an empirically more successful theory of gravitation but also because of its ability to satisfy general philosophical and methodological principles or preferences. His view, though set against the causal-explanatory theory of space-time, is still reminiscent of Einstein's view of Newtonian theory and special relativity. I will consider a number of philosophical and methodological principles that bear this out, and that seem to motivate the allegation of a conspiracy.

3.1. The action-reaction principle

The most notable of these philosophical principles is found in Einstein's view that inertial systems in Newtonian theory and special relativity are 'factitious causes' of inertial effects. This is exemplified in Einstein's (1916 [1952]) illustration involving two bodies S_1 and S_2 in relative rotation. S_1 is perfectly spherical while S_2 bulges at the

[16] Or, pressing the conspiracy metaphor further, do free particles conspire or is it the 'parts of space' that conspire to have certain symmetries?

equator in the manner of a body subject to a centrifugal force. Einstein asks for the explanation of the difference between these bodies, and he replies:

> Newtonian mechanics does not give a satisfactory answer to this question. It pronounces as follows: The laws of mechanics apply to the space R_1, in respect to which the body S_1 is at rest, but not to the space R_2, in respect to which body S_2 is at rest. But the privileged space R_1 of Galileo, thus introduced, is a merely *factitious* cause, and not a thing that can be observed. (Einstein, 1916 [1952], pp. 112-115)

Newtonian theory fails to give a satisfactory answer because the space R_1—namely, the inertial system—is invoked as the cause of the difference between the two bodies. This is philosophically objectionable because something unobservable is being granted a causal role and because this cause acts without being acted upon. Because of these features, Einstein holds the inertial system to be a factitious cause, which must be replaced by a genuine cause like the fixed stars. Einstein's illustration bears out what some have called 'the action-reaction principle': For something to be physical it cannot act without being acted upon.[17] This idea is also found in *Relativity: The Special and the General Theory* (1916 [1939], pp. 171-173), *The Meaning of Relativity* (1922 [1950], pp. 54-55), and 'On the aether' (1924 [1991], pp. 15-18).

Leaving aside the important fact that the action-reaction principle is based on a distortion of Mach's and especially Newton's views on rotation, the principle may be considered in its own right. I wish to consider three objections to it. To begin with, one might argue that the action-reaction principle is neither a metaphysical criterion of physicality nor an epistemological criterion of legitimate postulation but an arbitrary invention or 'mere hypothesis.' One might object, as Norton (1993, pp. 848-849) and Pitts (2006, p. 349) have, that a spurious necessity has been attributed to a principle that derives from Aristotelian and Leibnizian metaphysics, and thus is not empirically constrained. Second, one might argue that to think of the inertial system in Newtonian theory or special relativity as something that acts without being acted upon is to

---

[17]Brown (2005, pp. 140-142) points out that a similar principle can be distilled from Leibniz's philosophy: For something to be a substance it cannot act without being acted upon. See the *Discourse on Metaphysics* (section 14) and *Monadology* (proposition 61).

misunderstand its role in these theories. Neither Newtonian space-time nor the inertial system is being postulated as a theoretical entity that is the cause of inertial effects, for such a theoretical entity certainly would go against Newton's theory in which one is always dealing with interactions in which the participants enter reciprocally. Rather, Newtonian space-time is the structure that is *implicit* in Newton's account of causal influence; Minkowski space-time is the structure that is *implicit* in special relativity. To be sure, Newtonian space-time and Minkowski space-time express constraints on the possible evolution of fields—in just the same way that the Hilbert space structure of quantum mechanics and the configuration space of classical mechanics express constraints on possible states of systems. But such constraints do not represent the action of these structures on the fields in Einstein's sense.[18] Third, one might point out that the action-reaction principle amounts to a principle that excludes a priori the possibility that space-time is flat. But whether or not it makes sense to think of space-time as flat ought to be an empirical question. For example, the equivalence principle—it is impossible to distinguish locally between a homogeneous gravitational field and a uniformly accelerated frame—provides a basis for arguing that space-time is not flat. Therefore, there is certainly a basis for arguing that it does not make sense to think of space-time as unaffected by matter; but that argument is founded on an empirical hypothesis and not an a priori demand. Without the equivalence principle, one would have in Newtonian theory and special relativity space-time theories that are empirically unexceptionable. In such a case, the action-reaction principle would, strictly speaking, express nothing but a metatheoretical or metaphysical preference for a different sort of theory—a sort of theory that, in the absence of the equivalence principle or something like it, would be difficult to motivate empirically.

Brown also objects to the action-reaction principle. His objection stems, in part, from his accounts of Leibniz's and Newton's views on space and time: 'Nonentities do not act, so for Leibniz space and time can play no role in explaining the mystery of inertia.' (Brown, 2005, p. 142) Regarding Newton's view, he writes:

---

[18]The recognition of this important point can be found in DiSalle (2002, p. 182), Brown (2005, p. 139), and Pitts (2006, p. 349).

> For Newton, the existence of absolute space and time has to do with providing a structure, necessarily distinct from ponderable bodies and their relations, with respect to which it is possible systematically to define the basic *kinematical* properties of the motion of such bodies. For Newton, space and time are not substances in the sense that they can act, but are real things nonetheless. (Brown, 2005, p. 142)

But his main objection is that the non-dynamical affine and conformal structures whose application is controlled by Newtonian theory and special relativity do not figure in a causal explanation of inertial motion. They are 'a codification of certain key aspects of the behaviour of particles and fields.' (Brown, 2005, p. 142)

It is interesting to contrast Brown's view of inertial structure as a codification with Weyl's view. While Einstein held that inertial structure in Newtonian theory and special relativity is a factitious cause, Weyl held that once inertial structure is understood to be inseparable from gravitation it must be recognised as something that 'not only exerts effects upon matter but in turn suffers such effects.' (Weyl, 1927 [1949], p. 105). He referred to the inertial structure of space-time as 'the guiding field' in analogy with other physical fields, notably, fluids. Brown remarks that 'To appeal ... to the action of a background space-time connection in which the particles are immersed—to what Weyl called the "guiding field"—is arguably to enhance the mystery, not to remove it.' (Brown, 2005, p. 142) Weyl's account of the guiding field, with Brown's emphasis on its fluid aspect, enhances the mystery because free particles do not know what kind of space-time they are immersed in; they just do what they do.[19] Though Brown does not acknowledge Weyl's view except in these few remarks, it is safe to say that he dismisses the guiding field because it has a measure of explanatory power that is reminiscent of Einstein's causal-explanatory account and that goes against his view that space-time structure should be regarded as a codification or representational framework. His use of 'codification,' in fact, represents a deliberate deflation of inertial structure as something with explanatory power.

[19] But there is another reading of Weyl that focuses on his account of the 'world structure' that is exhibited in inertial motion rather than on his account of the guiding field. See, e.g., DiSalle (2006a, pp. 137-149; 2006b).

While Brown's objection to the action-reaction principle as an argument against Newtonian theory and special relativity is decisive, there remains the matter of the conspiracy allegation. The allegation appears to be bound up with the non-dynamical character of the global affine and conformal structures. It arises because free particles traverse geodesics of an affine structure that is fixed. So, though Brown rejects the action-reaction principle, something *close* to the principle appears to underlie the allegation: Newtonian theory and special relativity control the application of background-structures that shape the evolution of fields without being themselves shaped by them. In other words, there is no dynamical coupling of the affine and conformal structures to matter.

3.2. Global coordinate systems as an artifice of thought

Another philosophical intuition about inertial structure is found in Einstein's view that global coordinate systems are an artifice of our thought. Nature is indifferent to our choice of coordinate systems and does not single out certain kinds.[20] Einstein writes:

> What makes this situation appear particularly unpleasant is the fact that there should be infinitely many inertial systems, moving uniformly and without rotation with respect to one another, that are distinguished from all other rigid systems. (Einstein, 1951, pp. 27-29)

There are a number of objections to this intuition. First, as DiSalle (2002, pp. 178-180) has argued, saying that it is inherently mysterious that nature should distinguish certain kinds of inertial systems and their associated coordinate systems amounts to

[20]This intuition is bound up with the principle of general covariance, according to which the possible laws of physics should be restricted to those that admit a coordinate-independent formulation. The satisfaction of the principle of general covariance was supposed to be a philosophical advantage of Einsteinian gravitation, one that eliminated the 'epistemological defect' peculiar to Newtonian theory and special relativity with their global inertial frames. When Kretschmann showed in 1917 that Einsteinian gravitation is not unique in this respect, Einstein (1918 [2002], p. 242; 1951, p. 69) proposed an alternative to the principle that he took to capture the theory's characteristic feature and to surmount Kretschmann's objection: The possible laws of physics should not only admit coordinate-independent formulations but these formulations should be the simplest and most transparent ones available to them. Einstein claimed that this methodological principle has 'significant heuristic force.' The notion of 'theories that are not the simplest and most transparent in generally-covariant form' means 'theories that, in addition to being generally covariant, have other, non-trivial symmetries.' In this way, we are returned again to the a priori demand to eliminate theories that assert the possibility of global structure, theories whose status one would prefer to think of as an empirical question.

saying that it is inherently mysterious that space-time should have non-trivial symmetries. It may be that the existence of such symmetries and the dynamical laws that exhibit them are themselves mysterious, but the sense of mystery derives from a philosophical view, whatever that might be. Even if one were committed to such a view—one more akin to a form of apriorism than empiricism—it would be equally problematic that nature distinguishes conservative systems from all other physical systems. Second, one might point out that this intuition is bound up with a confusion about the relation between dynamical laws and coordinate systems; namely, the idea that Newton's laws only 'hold' in special coordinate systems. This idea can be found in the work of various authors (e.g., Einstein, 1951, p. 27; Cushing, 1998, p. 98), and there are passages in which Brown appears to be making such a claim: 'Inertial coordinate systems are those special coordinate systems relative to which the above conspiracy, involving rectilinear uniform motions, unfolds.' (Brown, 2005, p. 17) To put this another way, a class of special coordinate systems is being postulated in which the laws of motion—the laws that determine the alleged conspiracy—hold. But this is to put the cart before the horse. Newton's laws do not hold in special coordinate systems; they assert the *possibility* of coordinate systems in which all accelerations depend on impressed forces. The possibility of such systems is asserted by Newton's laws; it is not a prerequisite for them.[21]

Brown's view is strongly reminiscent of the Einsteinian intuition that global coordinate systems are an artifice of thought. For Brown, free particles conspire precisely because 'their motion appears extremely simple from the point of view of a family of privileged frames' (Brown, 2005, p. 142). The idea that this family of frames is privileged—and that there is something metaphysically objectionable about this or any other privileged family—is at the root of the conspiracy allegation.

3.3. Inertia as a concept in need of explanation

There is another idea underlying the allegation of a conspiracy that should be singled out. This is the idea that inertia in Newtonian theory and special relativity is in need of an explanation; in other words, that a 'dynamical origin of inertia' is required.

---

[21]See also DiSalle (2006a; 2002) in this regard.

This idea can be inferred from Brown's claim that inertia finds an explanation in Einsteinian gravitation because a geodesic theorem can be proved.

But the idea that there is something questionable about inertia because there is no more fundamental assumption from which it can be derived is based on a confusion. It is useful to recall how the concept of inertia arises in Newton's account of causal interaction. The laws of motion define and interpret the concepts of force and mass, and these concepts determine inertial motion as that state in which a mass is unacted upon by forces. In this way, inertia is tied to Newton's account of force; the concept of inertia cannot be articulated independently of dynamics. Neither inertia nor force, therefore, is a concept in need of an explanation. The laws of motion express mathematically formulated criteria for explicating and applying concepts that are *already* in use—e.g., the pre-theoretical concept of force as something determined by pushing, pulling or pounding some mass.

Even if Brown were to reject some or all of these ideas or to question the relative support that they lend to his total view, the burden is still on him to explain the philosophical and methodological basis that supports the allegation of a conspiracy—in other words, to explain why we should 'criminalise' perfectly good empirical behaviour. At the very least, it is safe to say that, though Brown may disagree with the precise contours of Einstein's view, he thinks that Newtonian theory and special relativity have a defect that is eliminated by Einsteinian gravitation.

4. If inertia is conspiratorial, should a broad class of theories be so characterised?

Leaving aside the philosophical and methodological basis that seems to motivate the allegation of a conspiracy, one might ask: If in fact there is something conspiratorial about inertial motion in Newtonian theory and special relativity, should a broad class of theories be so characterised? As a way of making sense of why one might read a conspiracy into the motion of free particles, one might suggest that, for a conspiracy theorist, *all* physical theories are conspiratorial to varying degrees. But I will argue that this suggestion, in a sweeping sense at least, deflects attention from the ideas that truly motivate the conspiracy allegation.

It is helpful to consider a few examples that do not bear out the correct sense of 'conspiracy.' For example, one might say: If there is something conspiratorial about inertia, it is remarkable that there is nothing conspiratorial about the conservation of linear momentum, according to which the total momentum in an isolated system is conserved.[22] Take, for example, a system of billiard balls in free space. The total momentum of the balls, before and after a collision, is conserved. One could impute to a conspiracy theorist the view that the balls conspire to interact only with each other and not with their environments, and to transfer momentum among themselves such that their total momentum is conserved.

Or, if inertia is a conspiracy, why is there nothing conspiratorial about the motion of non-interacting charged particles in electromagnetic fields? Every electron interacts with a given electromagnetic field in exactly the same way. One could suggest that a conspiracy theorist might say that the electrons conspire to act in this way. But the conspiracy theorist might reply that the field is a common cause to which their motion can be attributed, and so the alleged conspiracy is explained away.[23] On this line of reasoning, it is the presence of a field that charged particles can 'feel' that distinguishes their movement from the conspiratorial behaviour of free, uncharged particles.

To take another example, if inertia is a conspiracy, there is equal reason to think of equilibrium as arising from a conspiracy. There are many ways in which one might formulate such a conspiracy. To take a simple example, consider a rod in uniform translatory motion whose particles are in a stable equilibrium configuration. We might then Lorentz-boost the rod so it travels faster. The rod undergoes an acceleration for the

[22]One might think that the conservation of linear momentum is in fact an excellent example since, in the case of an isolated system, the principle of conservation of linear momentum simply *is* the principle of inertia.

[23]For the same reason, there is nothing conspiratorial about the motion of a pair of harmonic oscillators—of similar constitution that are isolated from, and thus unable to communicate with, each other—oscillating at the same frequency. Presumably, there is a common cause in their past for the synchrony of the forces that produce the oscillations.

duration of the boost before settling into a new stable equilibrium configuration. One could imagine that a conspiracy theorist might say that the particles conspire to reassemble themselves into the Lorentz-contracted rod. But no doubt a conspiracy theorist thinks of equilibrium as explicable by locally-acting forces and would therefore reject the comparison.

Though one might attempt to make sense of the conspiracy allegation in this way, this suggestion trivialises Brown's view. Not one of these examples captures the sense of 'conspiracy' or 'pre-established harmony' at issue for him. The principles and intuitions we have considered above reveal a view about absolute or global background-structures, structures that constrain the possible states of a system without themselves being influenced by the system's evolution.

To take an example that does seem to capture the correct sense of 'conspiracy,' we might look to the theory of weak interactions. Consider chiral or 'handed' processes, that is, processes whose theoretical account displays a left-right asymmetry. If inertia is a conspiracy, there is equally good reason for seeing something conspiratorial in the handedness exhibited by parity violation in the theory of weak interactions. Why isn't there anything conspiratorial about the decay of, e.g., cobalt 60 atoms? How do all of the cobalt atoms in the universe know to exhibit handedness in the same sense when they are oblivious to one another? One could say that they conspire to do so.[24] This example seems to have the salient feature: The relevant phenomenon, handedness, is tied to a global space-time structure, namely, orientation. The natural reply to the conspiracy theorist is, of course, that cobalt atoms display handedness not because they conspire but because parity-violating experiments are part of the evidentiary basis for orientation, one that we have specified in Minkowski space-time.[25] But, for a conspiracy theorist, the orientation of Minkowski space-time ought to be just as problematic as the global affine structure.

---

[24]This example is suggested by Brown (2005, p. 142; personal communication). A detailed discussion of parity violation in the decay of a cobalt 60 isotope, in the philosophical literature, can be found in Huggett (2000) and Pooley (2003).

[25]For details on specifying the orientability and orientation of a manifold, see Malament (2012, §2.1-2.2).

What seems to underlie Brown's view is the idea that absolute or global background-structures are what might be called 'unexplained foundations.' They are unexplained in the sense that they cannot be derived from more general assumptions. But the idea that there is a problem with unexplained foundations is itself problematic; it seems to reflect a foundationalism that is difficult to motivate empirically. Even if Brown acknowledges that 'all explanation must stop somewhere,' and so his view is not susceptible to any sort of regress, he still has to establish that, e.g., the global affine structures of Newtonian theory and special relativity are in need of an explanation—a view I have argued against in §3.3. Furthermore, if the notion of an unexplained foundation is indeed at the root of the conspiracy allegation, then there can hardly be much gain in explaining away the conspiracy of inertia by appealing to Einsteinian gravitation, for one can point to conspiratorial features even in that framework.

One could regard the global topology, metric signature, orientability, and temporal orientation, among other features of Einsteinian gravitation, as having the marks of a conspiracy, in this more specific sense.[26] Consider the Lorentzian signature of the pseudo-Riemannian metric. The Lorentzian signature of the metric does not come from the field equations; it is ensured by assuming the strong equivalence principle. Temporal orientation must also be specified. By examining these and other features of Einsteinian gravitation, we find that, while the affine and conformal structures are determined by the distribution of mass-energy, the theory requires the postulation of a number of quantities that do not come from the field equations alone. If there is any sense in which Newtonian theory and special relativity are conspiratorial, then certain features of Einsteinian gravitation can be said to be no less conspiratorial.

To such a challenge, a conspiracy theorist might reply that, whenever we can explain the conspiratorial features of a theory by showing how they emerge from dynamics at a lower level, we have improved our understanding to a certain degree, we have shown that something miraculous at one level has a deeper reason. For example,

---

[26] A discussion of the metric type (pseudo-Riemannian) and signature can be found in Brown (1997).

Einsteinian gravitation explains remarkable structural correspondences that were previously taken for granted. If indeed all physical theories are conspiratorial—in the specific sense of having unexplained foundations—such a strategy may be available to Brown. But this still fails to address the more important question of why we should regard the features in question as problematic. For example, though the Lorentzian metric signature, orientability, and temporal orientation in Einsteinian gravitation do not derive from the field equations, they are not brute posits; their application is controlled by empirical criteria. If the notion of an unexplained foundation is what is driving the conspiracy allegation, it is better by far to argue that there are no conspiracies at all.

5. The alleged explanation of inertia by Einsteinian gravitation

In this final section, I wish to address Brown's claim that inertial motion is explained by Einsteinian gravitation. I will begin by presenting that claim as well as Weatherall's challenge to it. I will then propose another way of thinking about the theorems that drive the claim.

Brown claims: 'GR is the first in a long line of dynamical theories ... that *explains* inertial motion.' (Brown, 2005, p. 141) Further on, he writes:

> Inertia, in GR, is just as much a consequence of the field equations as gravitational waves. For the first time since Aristotle introduced the fundamental distinction between natural and forced motions, inertial motion is part of the dynamics. It is no longer a miracle. (Brown, 2005, p. 163)

Brown's claim rests on the fact that a geodesic principle—free particles traverse timelike geodesics—can be derived from Einstein's field equations together with other assumptions.[27] The claim seems to presuppose a deductive-nomological scheme: One can take the field equations, energy and conservation conditions, and the resulting geodesic principle as *explanans*, then derive the motion of a free particle as *explanandum*.

[27]This is not quite Brown's claim. Brown claims that the geodesic principle follows *directly* from the field equations, a claim that Malament (2012), in light of the result of Geroch and Jang (1975), has shown to be not so straightforward. Brown (personal communication) grants this and maintains nonetheless that inertial motion is explained by Einsteinian gravitation.

There are various geodesic theorems, but Geroch and Jang's (1975) has a claim to being the most perspicuous, and I will limit my attention to it. Their theorem has the advantage of avoiding specific assumptions about the nature of the free, massive, test particle; it also has the advantage of showing that the particle traverses a curve in space-time rather than a line singularity. In any case, if any geodesic theorem can be said to figure in a deductive-nomological explanation of inertial motion, the Geroch-Jang theorem can be said to do so.

Brown's claim that inertial motion is explained by Einsteinian gravitation in a distinctive way was challenged by Weatherall (2011a), who showed that a geodesic principle can be derived in geometrised Newtonian gravitation. With this theorem in hand, Weatherall observes of inertial motion in geometrised Newtonian and Einsteinian gravitation: 'if *either* theory can be thought to explain inertial motion, then *both* do, in much the same way.' (Weatherall, 2011b, p. 280)

A line of objection is available to Brown: Both Geroch and Jang's and Weatherall's theorems proceed from the fundamental assumption of the conservation of energy-momentum: $\nabla_a T^{ab} = \mathbf{0}$. But, in Einsteinian gravitation, the conservation condition follows from Einstein's field equations; in geometrised Newtonian gravitation, it is an independent assumption. This line of objection is undermined by Weatherall, who argues that the conservation condition is a background assumption in both theories. It is an assumption that is more general than Newtonian and Einsteinian gravitation, an assumption about a general feature of the world that these theories and others respect.

I agree with Weatherall's observation that, if there is any sense in which Einsteinian gravitation can be said to explain inertial motion, then geometrised Newtonian gravitation can be said to explain it at least as well. But I will argue that the main contribution of these theorems lies not in their *explanation* but in their *explication* of inertial motion. By 'explication,' I mean the clarification afforded by these theorems of the conceptual structure of Einstein's theory—of a certain account of matter, of the assumptions required for describing the evolution of that matter, and of the interdependence of these conditions—rather than any importance they might have in some or another philosophical account of explanation.

The geodesic theorems make explicit an assumption that Newton makes in his own account of inertial motion. Current discussions of inertial motion in old-fashioned Newtonian theory focus on the laws of motion and the corollaries to the laws. But there is an underappreciated discussion in the Scholium to the Laws in which Newton shows that the third law—and thus, the conservation of momentum—is necessary for the first law to apply to systems that are subject to attractive forces. The passage of interest to us is Newton's demonstration of the third law of motion for attractions. The proof is straightforward. Take any two bodies *A* and *B* that attract each other. Place between them an obstacle that impedes their coming together. Suppose, for reductio, that *A* is more attracted to *B* than *B* is to *A*. That is, suppose that $F_{B \text{ on } A} \neq F_{A \text{ on } B}$. Bodies *A* and *B* will move towards each other, both eventually reaching the obstacle. The obstacle will be pressed more strongly by body *A* than by body *B*, and so will not remain in equilibrium between them. The stronger pressure of *A* against the system comprising the obstacle and *B* will make the entire system of the three touching bodies move straight forward in the direction from *A* to *B*. In empty space, the system will go on indefinitely with a motion that is always accelerated. But this contradicts Law 1. Hence, our supposition that $F_{B \text{ on } A} \neq F_{A \text{ on } B}$ must be false. Hence, $F_{B \text{ on } A} = F_{A \text{ on } B}$.[28]

Though this demonstration of the third law focuses on a system of bodies, it is significant that the law applies also to a single body that is itself a composite system. The principle of inertia, taken on its own, is satisfied only in the case of point-particles. For bodies that are themselves composed of particles, the third law is a necessary condition for inertial motion. That is, the system of particles making up a single body must interact in such a way that every force is balanced, failing which the body will accelerate by its own internal forces and violate the principle of inertia. Therefore, the notion of

[28]Newton is anticipating the application of the third law to the Solar System. He envisages the steps that he will take to show that the Solar System is effectively isolated.

equilibrium—as it pertains to a single, composite body as well as to a system of bodies—enters Newtonian theory only via the third law.[29]

Though it is often overlooked that the third law is a necessary condition for the first law to apply to systems held together by attractive forces, it was well understood by Newtonians in the eighteenth and nineteenth centuries with whom it was further elaborated and clarified. There are too many to consider individually, but it is important to mention d'Alembert, who, in his *Traité de Dynamique* (1743 [1967]), proposed a rational mechanics founded on laws of impact between perfectly hard bodies. Though d'Alembert was manifestly a Newtonian, the influence of Descartes on d'Alembert's thought can be clearly seen. D'Alembert sought to deduce the laws of mechanics from 'certain dispositions of size, figure and motion,' in other words, from a purely geometrical account. From such a clearly and distinctly known geometrical basis, his laws of motion would be necessary truths and his mechanics would be a genuine metaphysical discovery. This view led d'Alembert to propose laws of motion that are close to Newton's laws.[30] In spite of its Cartesian aspect, however, d'Alembert's mechanics is a restriction of Newtonian mechanics to the mechanics of rigid bodies.[31]

D'Alembert understood clearly that Newton's third law, and therefore the conservation of momentum, must be assumed to give an account of the transfer of motion from one body to another in a collision. Newton's third law enters d'Alembert's mechanics as 'the principle of equilibrium,' and the concept of equilibrium is the core of his 'general principle,' which we now know as 'd'Alembert's principle.'[32] The principle

---

[29]In Appendix A, I review how to formulate a principle of conservation of momentum for a Newtonian or special-relativistic system, with particular emphasis on how a state of equilibrium is obtained.
[30]It is worth noting that D'Alembert was reluctant to write of forces. He eschewed the lingering notion of inherent cause and the *vis viva* controversy. He insisted that 'force' is only that quantity with which we are acquainted through its *effects*.
[31]D'Alembert proposes to focus on bodies that act on one another by 'immediate impulse, as in the case of an ordinary impact' or by 'the interposition between them of some body to which they are attached' (d'Alembert, 1743 [1967], p. 49). He considers attractions to have been sufficiently well examined by Newton, and so sets these actions aside.
[32]D'Alembert's own statement of the principle (1743 [1967], p. 51) is not straightforward, but clear statements of its essential content can be found in the work of his successors. Thomson and Tait's statement is one such; other statements are found in Mach (1883 [1919], pp. 335-337). For a good, recent discussion of the principle, see Lanczos (1970).

asserts that 'all the forces acting on points of the system form, with the reactions against acceleration, an equilibrating set of forces on the whole system.' (Thomson and Tait, 1867 [1879], p. 248) This is the culmination of the *Traité de Dynamique*; it represents d'Alembert's attempt to reduce the laws of mechanics to a single principle.

With the general principle in hand, d'Alembert deduced three theorems. The first, which is of greatest interest to us, asserts that '[t]he state of motion or rest of the centre of gravity of several bodies does not change by the mutual action of these bodies among themselves, provided that the system is completely free' (d'Alembert, 1743 [1967], Part II, Ch. 2, Theorem I). In this way, the principle of the conservation of the centre of gravity is recovered from his general principle. He deduced a second theorem, according to which 'if weight or an accelerative force—constant for each body and different, if one wants, for each of them—acts on these bodies following parallel lines, the centre of gravity or rather the common centre of mass will describe the same curve that it would have described if these bodies had been free.' (d'Alembert, 1743 [1967], Part II, Ch. 2, Theorem II) This theorem generalises the first to encompass those situations in which an isolated system is acted upon by a force that is sufficiently distant for the system to be treated like an isolated or 'near enough' isolated system.[33] A third theorem generalises the first still further to encompass systems subject to a constraint. In the Scholia to the Theorems d'Alembert notes that these theorems are equally true for attractions; so, though he deliberately restricts his attention to rigid-body mechanics, he acknowledges that his principle has wider applicability. D'Alembert's laws of motion and his general principle establish clearly that the total 'quantity of motion' or 'momentum' in isolated systems of interacting bodies is conserved.

[33]Theorem II reveals d'Alembert's understanding of Newton's Corollary VI to the laws of motion. What is puzzling, however, is his suggestion that the forces may be different for each body. It may be that d'Alembert states Theorem II in the way that he does because he wants to acknowledge that Corollary VI contains an explicit (restrictive) hypothesis 'If bodies are ... urged by equal accelerative forces along parallel lines...' that is never strictly satisfied, except in the trivial case of zero accelerative forces. This reading seems to be supported by Sklar's (2013, pp. 120-121) interpretation of Theorem II as a generalisation of the principle of the conservation of the centre of gravity 'to include systems of particles all subject to the same external accelerating force, either constant and acting along parallel lines or directed to a point and distance-dependent.'

The centrality of the conservation principle to Newtonian theory was equally well understood by Thomson and Tait in their *Treatise on Natural Philosophy* (1867 [1879]). In their discussion of Newton's laws, they observe that

> Of late there has been a tendency to split the second law into two, called respectively the second and third, and to ignore the third entirely, though using it *directly* in every dynamical problem; but all who have done so have been forced *indirectly* to acknowledge the completeness of Newton's system, by introducing as an axiom what is called D'Alembert's principle, which is really Newton's rejected third law in another form. Newton's own interpretation of his third law directly points out not only D'Alembert's principle, but also the modern principles of Work and Energy. (Thomson and Tait, 1867 [1879], p. 240)

That the conservation of momentum is a necessary condition for the inertial motion of composite systems was noted in the same year by Maxwell in *Matter and Motion* (1867 [1888]):

> ... Newton goes on to point out the consequence of denying the truth of [the third law of motion]. For instance, if the attraction of any part of the earth, say a mountain, upon the remainder of the earth were greater or less than that of the remainer of the earth upon the mountain, there would be a residual force, acting upon the system of the earth and the mountain as a whole, which would cause it to move off, with an ever-increasing velocity, through infinite space. (Maxwell, 1876 [1888], p. 48)

This vivid illustration of the application of the third law to a body that is itself a composite system establishes in still another way the fundamental role of the conservation of momentum.

What we find in the work of D'Alembert, Thomson and Tait, Maxwell, and others is a deliberate attempt to give a perspicuous account of the necessity of the conservation of momentum for the inertial motion of composite systems, a relation that is manifest but not as prominent in Newton. As Geroch and Jang and Weatherall have shown, this relation is equally essential to Einsteinian gravitation and geometrised Newtonian gravitation. Therefore, in old-fashioned Newtonian theory no less than in geometrised theories the account of inertial motion is not determined by any single principle susceptible of separate explanation but by an interdependence of physical principles that must be assumed together. Old-fashioned Newtonian theory and the geometrised theories are in fact strongly analogous in their accounts of inertial motion: The third law of motion is to old-fashioned Newtonian theory as the conservation principles are to the geometrised theories. This analogy is clearly exhibited by the geodesic theorems, and it highlights the sense in which their contribution to our understanding does not lie in their explanation of inertial motion but in their explication of it.

The sense of 'explication' in question has nothing to do with our ability to derive a previously unprovable proposition from a new theory, though, in the cases that concern us, the proofs contribute to that explication. Nor does this sense of explication have anything to do with any particular philosophical account of scientific explanation; and so it is independent of the success or failure that attaches to such an account. Rather, the explication is the fruit of an *analysis* that began with the question, on what assumptions does our use of the concept of inertia depend? The analysis reveals that, in both old-fashioned Newtonian theory and in geometrised theories, inertia depends fundamentally on the conservation of momentum. Far from a concern with explaining the causal or dynamical origin of inertia, the geodesic theorems explicate the concept by revealing the connections between inertia and other concepts.

6. Conclusion

I set out to evaluate Brown's account of inertial motion in Newtonian theory and special relativity; in particular, his claim that there is something objectionable—something conspiratorial—about inertia in these theories. I presented and clarified the conspiracy allegation, and I argued that it is motivated by a commitment to a number of philosophical and methodological principles or intuitions that are reminiscent of Einstein's view; namely, the action-reaction principle, the idea that global coordinate systems are an artifice of thought, and the idea that inertia in Newton's theory is in need of an explanation. These principles reveal that the conspiracy allegation is bound up with a view according to which there is something problematic about absolute or global space-time structures. I argued that, even if Brown does not accept some or all of these principles, the onus is still on him to explain why there is anything problematic about inertial motion in Newtonian theory and special relativity.

I then asked, if there is something conspiratorial about inertia, should a broad class of theories be so characterised? I considered the seemingly natural suggestion that, for a conspiracy theorist, all physical theories are conspiratorial. I examined and rejected a sweeping sense of 'conspiracy' that trivialises Brown's view. I then examined a narrower sense that is bound up with the notion that absolute or global background-structures are an unexplained foundation, and I pointed out that Einsteinian gravitation also has such features. I argued that, if indeed the conspiracy allegation is driven by this idea, then it is better to argue that there are no conspiracies at all.

Last, I addressed Brown's claim that inertia is explained by Einsteinian gravitation because a geodesic principle can be derived from the field equations. I reviewed Weatherall's (2011b) challenge to Brown's claim. Weatherall argued that, if there is any sense in which Einsteinian gravitation can be said to explain to inertia, then geometrised Newtonian gravitation explains it at least as well. While I agreed with Weatherall, I argued that there is a better way of thinking about the geodesic theorems. That is, their main contribution lies not in their explanation of inertial motion but in their explication of it. This explication is independent of any philosophical account of explanation under which inertia can be subsumed; it is concerned with clearly exhibiting the assumptions on which our use of the concept depends.

I argued that the geodesic theorems of Geroch and Jang (1975) and Weatherall (2011a) explicate inertial motion by making perspicuous the dependency of inertial motion on the conservation of momentum. This is manifest, though under-appreciated, in Newton's own account of inertia, and I argued that the work of his successors—notably, d'Alembert, Thompson and Tait, and Maxwell—represents a deliberate attempt to establish the fundamental importance of the conservation principle. In spite of their important differences, old-fashioned Newtonian theory, geometrised Newtonian gravitation, and Einsteinian gravitation are strongly analogous in their accounts of inertial motion.

References

Brown, H. (2005). *Physical Relativity: Space-time Structure from a Dynamical Perspective.* Oxford: Oxford University Press.

Brown, H. (1997). On the role of special relativity in general relativity. *International Studies in the Philosophy of Science, 11*, 67-81.

Cushing, J. (1998). *Philosophical Concepts in Physics*. Cambridge: Cambridge University Press.

D'Alembert, J. (1967). *Traité de dynamique*. Brussels: Impression Anastaltique. (Original work published in 1743)

Demopoulos, W. (2000). On the origin and status of our conception of number. *Notre Dame Journal of Formal Logic*, 41, 210-226.

DiSalle, R. (2012). Analysis and Interpretation in the Philosophy of Modern Physics. In M. Frappier, D. Brown, & R. DiSalle (Eds.), *Analysis and Interpretation in the Exact Sciences: Essays in Honour of William Demopoulos* (pp. 167-191). Dordrecht: Springer.

DiSalle, R. (2006a). *Understanding Space-time.* Cambridge: Cambridge University Press.

DiSalle, R. (2006b). Mathematical Structure, 'World Structure,' and the Philosophical Turning-Point in Modern Physics. In V. F. Hendricks et al., *Interactions: Mathematics, Physics and Philosophy, 1860-1930* (pp. 207-230). Dordrecht: Springer.

DiSalle, R. (2002). Reconsidering Ernst Mach on Space, Time, and Motion. In D. Malament (Ed.), *Reading Natural Philosophy* (pp. 1-18). Chicago: Open Court.

DiSalle, R. (1995). Spacetime Theory as Physical Geometry. *Erkenntnis, 42*, 317-337.

Einstein, A. (2002). Prinzipielles zur allgemeinen Relativitätstheorie. In M. Janssen, R. Schulmann et al. (Eds.), *The Collected Papers of Albert Einstein, vol. 7* (pp. 38-41). Princeton: Princeton University Press. (Original work published in 1918)

Einstein, A. (1991). On the Ether (S. Saunders, Trans.). In S. Saunders & H. Brown (Eds.), *Philosophy of Vacuum* (pp. 13-20). Oxford: Oxford University Press. (Original work published in 1924)

Einstein, A. (1952). The Foundation of the General Theory of Relativity. In H. A. Lorentz, A. Einstein, et al., *The Principle of Relativity* (pp. 111-164). (Original work published in 1916)

Einstein, A. (1951). Autobiographical Notes. In P. Schilpp (Ed.), *Albert Einstein: Philosopher-Scientist* (pp. 1-95). New York: Tudor Publishing Company.

Einstein, A. (1950). *The Meaning of Relativity*. London: Methuen and Co. (Original work published in 1922)

Einstein, A. (1939). *Relativity: The Special and the General Theory*. London: Methuen & Co. (Original work published in 1916)

Geroch, R. & P.-S. Jang. (1975). Motion of a body in general relativity. *Journal of Mathematical Physics*, 16, 65–67.

Hagar, A. (2008). Length Matters: The Einstein-Swann Correspondence. *Studies in the History and Philosophy of Modern Physics*, 39(3): 532-556.

Huggett, N. (2000). Reflections on Parity Nonconservation. *Philosophy of Science, 67*, 219-241.

Janssen, M. (2009). Drawing the Line between Kinematics and Dynamics in Special Relativity. *Studies in History and Philosophy of Modern Physics*, 40, 26-52.

Lanczos, C. (1970). *The Variational Principles of Mechanics*. Toronto: University of Toronto Press.

Lange, L. (1885). Ueber das Beharrungsgesetz. *Berichte der Koniglichen Sachsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-physische Classe,* 37, 333-351.

Leibniz, G. (1970). *Philosophical Papers and Letters* (L. Loemker, Ed.). Chicago: University of Chicago Press.

Mach, E. (1919). *The Science of Mechanics: A Critical and Historical Account of its Development* (T. McCormack, Trans.). Chicago: Open Court. (Original work published in 1883)

Malament, D. (2012). A Remark About the ‘Geodesic Principle’ in General Relativity. In M. Frappier, D. Brown, & R. DiSalle (Eds.), *Analysis and Interpretation in the Exact Sciences: Essays in Honour of William Demopoulos, vol. 78* (pp. 245-252). Dordrecht: Springer.

Malament, D. (2012). *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory*. Chicago: University of Chicago Press.

Maxwell, J. C. (1888). *Matter and Motion*. London: Society for Promoting Christian Knowledge. (Original work published in 1867)

Nerlich, G. (1976). *The Shape of Space*. Cambridge: Cambridge University Press.

Newton, I. (1999). *The Principia: Mathematical Principles of Natural Philosophy* (I. Cohen & A. Whitman, Trans.). Berkeley: University of California Press. (Original work published 1726)

Norton, J. (2008). Why Constructive Relativity Fails. *British Journal for the Philosophy of Science*, 59(4): 821-834.

Norton, J. (1993). General covariance and the foundations of general relativity: Eight decades of dispute. *Reports of Progress in Physics*, 56, 791-861.

Pfister, H. (2004). Newton's first law revisited. *Foundations of Physics Letters, 17*, 49-64.

Pitts, B. (2006). Absolute Objects and Counterexamples. *Studies in History and Philosophy of Modern Physics, 37*, 347-371.

Pooley, O. (2003). Handedness, parity violation, and the reality of space. In K. Brading & E. Castellani (Eds.), *Symmetries in Physics*. Cambridge: Cambridge University Press.

Sklar, L. (2013). *Philosophy and the Foundations of Dynamics*. Cambridge: Cambridge University Press.

Thomson, W. & P. G. Tait. (1879). *Treatise on Natural Philosophy, vol. I, Part I. New Edition.* Cambridge: Cambridge University Press. (Original work published in 1867)

Weatherall, J. (2011a). Motion of a body in Newtonian theories. *Journal of Mathematical Physics*, 52, 1-16.

Weatherall, J. (2011b). On the status of the geodesic principle in Newtonian and relativistic physics. *Studies in History and Philosophy of Modern Physics,* 42, 276-281.

Weyl, H. (1949). *Philosophy of Mathematics and Natural Science* (O. Helmer, Trans.). Princeton: Princeton University Press.

Chapter 3

Friedman's Thesis[†]

1. Introduction

In 'Two dogmas of empiricism' (1951), Quine represented scientific knowledge as a web of belief in which no satisfactory analytic-synthetic distinction can be drawn. In the absence of a suitably broad notion of analyticity, no propositions deserve to be singled out as being true in virtue of their meanings or as having any other measure of necessity, apriority or epistemic security. Quine acknowledged that certain stipulations like definitions are undoubtedly analytic, but that we can have no assurance that the propositions of mathematics are epistemologically distinguished from physical propositions just because they have been stipulated to be analytic. The arbitrariness that attaches to any such stipulation led him to reject the analytic-synthetic distinction.[34]

This view, while motivated by a particular understanding of the logical empiricists' approach to the analysis of theories, led Quine to the far more general view that no distinctions of kind can be drawn among the propositions comprising our web of belief. There is no distinction of kind between mathematical and physical propositions, and no distinction between these propositions and philosophical propositions. Philosophy is not a form of meta-theoretical or transcendental analysis, as has long been maintained; rather, philosophy is itself a part of scientific enquiry. Quine called this view 'naturalism.'

Friedman's view is set against this naturalism. Friedman sees in the conceptual structures of Newtonian and Einsteinian gravitation a clear basis for correcting Quine. He defends the idea that a framework of physical knowledge is stratified, and he argues that, among the different kinds of principles comprising a theory, there are certain principles—'constitutive principles'—whose identification is indispensable to a satisfactory

[†]A version of this chapter is under review at the time of submission.

[34]Demopoulos (2013) proposes a way of establishing some of the principal conclusions that Carnap based on the analytic-synthetic distinction and that he defended in his long-standing controversy with Quine. It is significant that this proposal does not trade on the notions of truth in virtue of meaning, convention, or stipulation.

methodological analysis of physics. Friedman's proposal culminates in the thesis that I call *Friedman's thesis*: Revolutionary theory change proceeds by deliberate philosophical reflection on constitutive principles.[35] My goal in this essay is to explicate and evaluate Friedman's thesis.

2. Friedman on the structure of physical theories

Friedman defends an account of the structure of theories and theory change in which there are three levels of enquiry. The *first level* is comprised of principles that are epistemologically distinguished by the fact that they determine the framework of investigation; that is, they articulate a network of theoretical concepts and their physical interpretations. The *second level* is comprised of empirical hypotheses that are formulable within that framework. The *third level* is comprised by distinctly philosophical or meta-theoretical principles that motivate discussions of the framework-defining principles and the transition from one theory to another.

Those principles that determine the framework of investigation Friedman calls 'constitutive principles.' Those that Friedman calls 'mathematical principles' supply the formal background or language that makes it possible to articulate a theory's basic concepts and that makes particular kinds of applications possible. We find, among other examples, the calculus, linear algebra, and Riemann's theory of manifolds. But there are other constitutive principles that have a more complex character: These 'coordinating principles' interpret the concepts that are necessary for physics as we understand it; they express criteria by which concepts such as force, mass, motion, electric field, magnetic field, space, and time may be applied. The mathematical principles are important, but the coordinating principles control the application of the mathematics, something the mathematics by itself does not do.

[35]The thesis in question is exemplified in *Dynamics of Reason* (2001), and aspects of it are developed in 'Carnap and Quine: Twentieth-century echoes of Kant and Hume' (2006), 'Synthetic history reconsidered' (2010a), 'A post-Kuhnian approach to the history and philosophy of science' (2010b), and 'Einstein and the a priori' (2011).

Friedman's most careful characterisation of a constitutive principle is found in the following passage of *Dynamics of Reason*: 'What characterizes the distinguished elements of our theories is ... their special *constitutive function*: the function of making the precise mathematical formulation and empirical application of the theories in question first possible.' (Friedman, 2001, p. 40) This characterisation shows clearly that for Friedman there are two kinds of constitutive principles: those that supply the mathematical language that makes it possible to formulate the theory and that makes certain applications possible, and those that have a coordinating function.

The notion of coordination peculiar to Friedman's characterisation has its origin in Reichenbach (1920 [1965], section V; 1924 [1969], §2; 1928 [1958], §4). Reichenbach proposed an account of the structure of theories in which he defended a special class of physical principles that he called 'coordinative definitions.' These principles interpret theoretical concepts by associating them with something in the world of experience. To take what is perhaps the simplest example, Euclidean geometry becomes a theory of *physical geometry* by means of two coordinative definitions: The principle 'light rays may be treated as straight lines' interprets the Euclidean concept of straightness; the principle 'practically rigid bodies undergo free motions without change of shape or dimension' interprets the concept of congruence. Since the possibility of carrying out Euclidean constructions implicitly presupposes the concepts of straightness and congruence, these principles control the application of Euclidean geometry.[36] Because of this interpretive function, Reichenbach regarded coordinative definitions as relativised but nonetheless 'constitutive a priori principles' that serve to apply an uninterpreted corpus of mathematics—the 'categories'—to the world of experience. But they are not true absolutely; they develop along with physical theory, and so are relativised to particular contexts of enquiry. Furthermore, there exist coordinations that cannot be held at the same time. Reichenbach took it as a sort of Kantian principle that coordination is arbitrary, in the sense that no facts can fail to be accommodated within the framework of

[36]See Appendix B for a discussion of these principles and their identification by Helmholtz and Poincaré. The appendix also contains a brief discussion of Maxwell's equations.

a priori principles. But experience can show that combinations of individually reasonable coordinations can be inconsistent.

Carnap (1934 [1951], p. 78) initially accepted Reichenbach's notion of a coordinative definition without modification. But, in subsequent work, he came to regard that notion as an oversimplification. Where Reichenbach understood coordinative definitions to give a direct and complete interpretation of theoretical terms in terms of our observational vocabulary, Carnap held that such principles, which he came to call 'correspondence rules,' interpret them only indirectly, and so partially and incompletely. In a mathematical theory, a theoretical term like 'number' can be interpreted completely in logical terms. But this is not possible in the case of modern physical theories. Given a theory of modern physics, in which one takes as primitive those theoretical terms that figure in a few fundamental laws of great generality, the correspondence rules 'have no direct relation to the primitive terms of the system but refer to terms introduced by long chains of definitions .... For the more abstract terms, the rules determine only an *indirect interpretation*, which is ... incomplete in a certain sense.' (1939, p. 65) The same view is found in 'The Methodological Character of Theoretical Concepts' (1956) and in *Philosophical Foundations of Physics* (1966).

The oversimplification that Carnap identified in Reichenbach's account is avoided by Friedman's characterisation of a coordinating principle. But Friedman's notion of a constitutive principle is broader than Reichenbach's notion of a coordinative definition and Carnap's notion of a correspondence rule; it encompasses principles that have a coordinating function, like Reichenbach's and Carnap's principles, as well as mathematical principles. What is common to Reichenbach, Carnap, and Friedman is the view that frameworks of physical knowledge are *stratified*. Those principles that are constitutive of the objects of scientific knowledge are not of the same kind as properly empirical hypotheses since they make those hypotheses 'possible.'

This account of the structure of theories stands in sharp contrast with Quine's 'naturalism,' according to which no elements of the web of belief have any distinguished epistemological status. Quine regarded set theory—and therefore all of mathematics—as continuous with physics. Philosophy, as a chapter of psychology, is part of this continuum. With this naturalism, it is precisely the stratification characteristic of the logical empiricists' approach that is lost. For Friedman, Quine's account of the structure of theories is a failure: It does not recognise that distinguishing constitutive principles from empirical hypotheses is essential to a satisfactory methodological analysis of physics, and it fails to appreciate the role played by constitutive principles in the articulation of basic theoretical concepts. This is what is lost with the replacement of *stratification* with the relative *centrality* of certain propositions in our web of belief. This, for Friedman, is the real divergence between Quine and the logical empiricists, Carnap foremost among them.

What is more, Friedman claims that careful attention to the history of physics shows that revolutionary theory change proceeds by deliberate philosophical reflection on constitutive principles. Friedman offers this proposal as an alternative to Kuhn's characterisation of revolutionary theory change as the result of a paradigm shift. The proposal is intended to illuminate revolutionary theory change not only in space-time physics but in physics in general. I will return to Friedman's account of theory change in §5.

Friedman's work to restore a proper understanding of the stratification of our frameworks of physical knowledge is a significant contribution to methodology. But his characterisation of a constitutive principle nonetheless includes too much of a theory's formal apparatus in the category of its constitutive principles. Friedman's inclusion of mathematical principles is motivated by his view that the role of mathematics in physics is distorted when it is regarded as just another element of our web of belief.[37] I agree with Friedman about this, but, in what follows, I will argue that 'constitutive principle' has a narrower reference. I will argue that mathematical principles should be distinguished

[37]This view is found in *Dynamics of Reason* (2001) and also in 'Carnap and Quine: Twentieth-century echoes of Kant and Hume' (2006). Friedman argues that physical theories presuppose a number of mathematical theories for the articulation of basic concepts and for the generation of empirical predictions. But it is also essential to his view that constitutive principles—both mathematical and coordinating principles—define a space of intellectual possibilities. See Friedman (2001, p. 84).

from coordinating principles. Both mathematical principles and coordinating principles are non-factual, but for different reasons. Coordinating principles are 'answerable' to the world; mathematical principles are not. The former fix an interpretation of the world; the latter, as part of the formal background or language, are prerequisites to this. To put this another way, mathematical principles and coordinating principles have different criteria of truth. This is not to diminish the importance of the mathematical principles, but to emphasise that only the coordinating principles are constitutive—in the sense that they interpret theoretical concepts by expressing criteria for their application.

3. The constitutive basis of Newtonian gravitation

Friedman brings his approach to the analysis of physical theories to bear on Newtonian and Einsteinian gravitation. I will briefly present Friedman's analysis of Newtonian gravitation, extending and sharpening a few important points. I will also consider Friedman's claim that the calculus and Euclidean geometry are constitutive presuppositions.

In keeping with the approach presented above, Friedman distinguishes constitutive principles, both mathematical and physical, and the framework they determine from empirical hypotheses whose formulation that framework permits. He presents Euclidean geometry, the calculus, and the laws of motion as constitutive presuppositions of the law of universal gravitation, which is a genuine empirical hypothesis (2001, Lecture II; 2010a, pp. 696-729; 2010b, p. 500; 2011, p. 1).

Friedman asks us to consider the relation between the law of universal gravitation and the laws of motion. The law of universal gravitation asserts that every object in the universe attracts every other object with a force that is directed along the line intersecting the two objects, and that is proportional to the product of their masses and inversely proportional to the square of the distance between them. The concepts of mass and force to which the law refers, however, are constituted by the second law of motion, a law that itself presupposes a state of inertial motion. And that state of motion, in turn, is constituted by the first and third laws. Only when they are taken together do the laws of motion constitute the concepts of force, mass, and inertial motion. These concepts have intuitive, pre-systematic meanings that are independent of the laws; for example, one may speak of a push-force or tension-force. But, while such meanings may suffice for everyday purposes, they provide no basis for recognising an instance of the concept in an unambiguous and intersubjective manner, and, most importantly, they provide no basis for measuring force. It is the laws of motion that constitute the concepts of force, mass, and inertial motion by expressing criteria for their application. The sense of 'constitutive' at issue is not merely that the laws of motion define the concepts to which they refer but that they interpret them. That is to say, they associate theoretical concepts with empirically measurable correlates.

What is more, the laws of motion are constitutive not only of a particular conception of force, mass, and inertial motion but of a particular conception of space, time, and causality.[38] With the development of a more abstract view of geometry in the twentieth century, it was shown that the space-time structure determined by the laws can be treated as a particular kind of four-dimensional affine space, with a specific foliation, and with temporal and spatial metrics having certain properties.[39] But this affine space, taken in itself, is just an instance of an abstract geometry of the sort made possible by twentieth-century methods. It is the laws of motion that control its application in physical theory.

All three laws of motion taken together therefore are constitutive of the framework that the inverse-square law requires for its formulation: They determine and control the application of the framework of empirical investigation—a framework that allows us to pose questions to be answered by the phenomena of motion, including questions about central forces to which the inverse-square law is an answer. Our ability to pose questions depends foremost on the conception of causal explanation that is

[38]It is noteworthy that, for Kant, the employment of our metaphysical concepts of causal interaction, force, motion, space, and time is inseparable from Newton's laws.

[39]The invariance of the velocity of light is another basis on which to treat space-time as a different kind of affine space.

expressed in the framework of the laws. The framework identifies the sorts of changes that are objectively measurable and that are indicative of the action of some cause.

Having addressed the relation between the law of universal gravitation and the laws of motion, Friedman asks us to consider next the relation between the laws of motion and the calculus. The concept of acceleration that figures in the second law is a quantity that requires the notion of instantaneous rate of change: Acceleration is the instantaneous rate of change of velocity, which is itself the instantaneous rate of change of position. Only with the calculus do we have an account of limiting processes and instantaneous rates of change; in short, a mathematical account of continuity. Friedman claims that the calculus, therefore, is a constitutive presupposition of Newtonian dynamics.

But, contrary to Friedman's view, the calculus should be characterised as part of the formal background or language that makes possible particular applications of Newton's laws and not as part of the theory's constitutive basis.[40] This is not to say that the calculus is not necessary for formulating Newtonian theory. One might say that the account of force in the second law is intelligible without the calculus; for example, one might suggest that force can be understood as the instantaneous result of pulling, pushing or pounding some mass. But it is the calculus that allows us to formulate the notion of a continuously-varying power, to develop the idea, for example, that Keplerian motion might be the manifestation of a yet-undetermined but continuously-varying force.[41] In this respect, the calculus certainly makes possible particular applications of the laws of motion, but it is not constitutive; it does not itself have an interpretive function. Further principles are required for its application in physical theory.

[40]More generally, the calculus is part of the theory's inferential apparatus: It tells us how particular quantities evolve given some initial data.

[41]By way of another example, one could also say that the Galilean composition of motions can be understood without the calculus; for example, the composition of the Earth's annual revolution around the sun and its diurnal rotation. But it is the calculus that allows us to treat arbitrary, continuous orbits as instances of the Galilean composition of motions.

For the same reason, Euclidean geometry should also be characterised as part of the formal background that Newtonian gravitation presupposes for its formulation. For Friedman, as for Kant, Euclidean geometry is a constitutive presupposition of Newtonian gravitation. Newton's own development of his theory presupposes the straightedge-and-compass constructions of Euclidean geometry, which Kant took to reflect our spatio-temporal intuition. As we have noted, however, the space-time structure of Newton's completed dynamical theory is a particular kind of four-dimensional affine space, with separate spatial and temporal metrics. The laws of motion, therefore, control the application of *this* particular affine space and *not* the framework of Euclidean geometry whose interpretation the laws already take for granted. But, setting this aside, what Friedman's claim most clearly brings to light is the sense of 'constitutive presupposition' at issue for him. Here the sense expressed is close to the ordinary dictionary sense of 'presupposition,' namely, 'a thing tacitly assumed beforehand at the beginning of a line of argument' (*OED*). This sense of 'constitutive presupposition' can also be found in the work of Poincaré, who pointed out that a geometry must be presupposed for the construction of a dynamical theory, but that doing so neither assumes nor precludes the possibility that the completed theory or another theory that is in some sense more fundamental may lead us to revise our presuppositions about geometry. Such a sense may be defensible, but it is different from the one exemplifed in the laws of motion.

Besides the importance of distinguishing principles that are answerable to the world from those that are not, another important result of separating principles that interpret theoretical concepts from mathematical principles or prerequisites is that it defends the idea of a constitutive principle against trivialisation. It might be argued that what is constitutive is relative to a theory's particular axiomatisation or formalisation, and, since what is constitutive in one axiomatisation or formalisation of a theory is not constitutive in another, the very idea of a constitutive principle is a wash. The limitation I propose permits agreement on the principles that interpret the theoretical concepts of a given theory, even if that theory admits of an alternative axiomatisation or formalisation. Newtonian theory, for example, admits of various mathematical settings, and those mathematical settings peculiar to analytic mechanics are radically different from

Newton's own constructive methods. But, even in the Lagrangian formulation, for example, Newtonian theory still expresses the same fundamental picture of space, time, and causality. However it is developed mathematically, Newtonian theory is the theory whose basic theoretical concepts are constituted by the laws of motion.

In spite of these criticisms of Friedman's characterisation of the calculus and Euclidean geometry, and so of the scope of his characterisation of a constitutive principle, his criticism of Quine's naturalism remains intact. His approach to the analysis of physical theories aims to clarify the relations between the inverse-square law, the laws of motion, the calculus, and Euclidean geometry. And that analysis does succeed in showing that these parts of the total framework of Newtonian gravitation are not of the same kind.

Much as this analysis is clarifying, there is a further sense in which the laws of motion are constitutive, one essential to Newton's own understanding of gravitation. One can read Newton's argument from the framework determined by the laws of motion to his gravitation theory as arising from a question about the applicability and adequacy of that framework for giving an account of celestial motion. By pressing the laws of motion as far as they can be pressed, that is, by boldly postulating that *all* bodies influence each other as per the third law of motion, we are driven to the hypothesis that there is an attraction—a 'universal gravitation'—between all bodies that acts instantaneously at a distance. It is only with this empirical hypothesis that an estimate of the masses of the bodies comprising a planetary system becomes possible; that an estimation of the centre of mass of the system is possible; and only with this hypothesis, therefore, that a planetary system can be considered as approximating an inertial frame. The form of the gravitational interaction, however, is not postulated but 'deduced from the phenomena' of planetary motion and gravitational free fall once these phenomena are understood within the necessary and sufficient framework of the laws of motion. Furthermore, not only was universal gravitation an open question but so also was the characterisation of its properties; for example, does gravitation propagate through a medium or immediately at a distance?

The idea that, given the framework of the laws of motion, an account of celestial motion is an open empirical question was central also to Euler's understanding of Newton.[42] Euler (1775 [1768-1772]), for all his work to turn Newton's theory into what we now recognise as 'Newtonian mechanics,' rejected action at a distance. He hoped a viable vortex theory would replace Newton's theory of attraction. But, in spite of that, he recognised the difference between the parts of Newton's theory that *any* theory of motion must constitutively presuppose and hypotheses formulable within that framework:

> Euler saw the difference between the elements of Newton's theory that were, so to speak, idiosyncratically Newtonian—above all the idea that univeral gravitation is the sole force at work in the Solar System—and those that represented the common basis of all work in mechanics as then understood, especially the laws of motion and their underlying framework of space and time. Thus he acknowledged the distinction between the physical hypotheses that one might prefer, pursue, and evaluate within the general framework of mechanics, and the conceptual framework without which such hypotheses could not even be intelligible. (DiSalle, 2006a, p. 51)

Euler recognised clearly that the laws of motion constitute a framework of investigation that is independent of hypotheses about what sorts of forces there are. He allowed for the possibility of an alternative to universal gravitation, all the while recognising that a Cartesian or any other proponent of a vortex hypothesis must himself presuppose the laws of motion in giving that alternative.[43]

4. Friedman's analysis of Einsteinian gravitation

Let us turn now to Friedman's analysis of Einsteinian gravitation. Friedman regards that framework as the outcome of three revolutionary advances, namely, the development of Riemann's theory of manifolds, Einstein's insight of 1907 that is summarised in the equivalence principle, and Einstein's field equations. All three were brought together to eliminate the contradiction between the instantaneous action at a distance postulated by Newtonian gravitation and the invariance of the velocity of light in

---

[42]See Aiton (1972) and especially Wilson (1992) for discussions of Euler's rejection of action at a distance and for further references.

[43]It is noteworthy that the laws of motion are implicitly presupposed not only in Cartesian physics but also in the work of Galileo, Huyghens, Wallis, and Wren on projectile motion and elastic collisions. This was Newton's argument for taking them to be axioms.

special relativity. In keeping with his approach to the analysis of theories, Friedman distinguishes constitutive principles, both mathematical and physical, from properly empirical hypotheses:

> ... the three advances together comprising Einstein's revolutionary theory should not be viewed as symmetrically functioning elements of a larger conjunction: the first two [Riemann's theory and the equivalence principle] function rather as necessary parts of the language or conceptual framework within which the third [the field equations] makes both mathematical and empirical sense. (Friedman, 2001, p. 39)

Further on, we find a sharper statement of Friedman's view that the equivalence principle functions to coordinate Einstein's field equations with experience:

> Einstein's field equations describe the variations in curvature of space-time geometry as a function of the distribution of mass and energy. Such a variably curved space-time structure would have no empirical meaning or application, however, if we had not first singled out some empirically given phenomena as counterparts of its fundamental geometrical notions—here the notion of geodesic or straightest possible path. The principle of equivalence does precisely this, however, and without this principle the intricate space-time geometry described by Einstein's field equations would not even be empirically false, but rather an empty mathematical formalism with no empirical application at all. (Friedman, 2001, pp. 38-39)

This is the core of Friedman's analysis of Einsteinian gravitation: Riemann's theory of manifolds and the equivalence principle are constitutive presuppositions of Einstein's field equations.

But there is a further aspect to this analysis that I will touch on only briefly: Friedman claims that the equivalence principle is elevated to the status of a definition in Poincaré's sense: 'In using the principle of equivalence to define a new four-dimensional inertial-kinematical structure, therefore, Einstein has 'elevated' this merely empirical fact to the status of a "convention or definition in disguise"' (Friedman, 2011, p. 8).[44] This claim is motivated by the fact that, though both Newton and Einstein were aware that

[44] In some passages, such as the one quoted, Friedman writes 'elevated to the status of a definition'; in others (e.g., Friedman, 2001, pp. 90-91), he writes 'elevated to a coordinating principle.' Whichever term is used, these passages reveal a central feature of Friedman's view, namely, that taking some principle as a new constitutive principle involves an element of decision or convention.

inertial mass and gravitational mass are indistinguishable, only Einstein took that indistinguishability as a basis for reinterpreting the concept of inertial motion. This claim is analogous to Friedman's claim that Einstein elevated the light postulate to the status of a definition: Whereas Lorentz took the invariance of the velocity of light as something to be explained by his theory of the electron, Einstein elevated it to the status of a definition, which he took as a basis for reinterpreting the concept of simultaneity.[45]

I will take issue with Friedman's claim that the equivalence principle and Riemann's theory are constitutive presuppositions by recalling Einstein's argument for a new concept of inertial motion and by contrasting that argument with Friedman's account. In my presentation of Einstein's argument I have not hesitated to make use of conceptual and mathematical insights that were gained only later. This departure from the actual history focuses attention on the shape of the argument without getting tangled up in questions about the success of individual steps.

4.1. The argument for curvature

Einstein took the first steps towards the inertial frame concept characteristic of his gravitation theory in 1905. The 1905 inertial frame concept emerged as the result of Einstein's recognition that the nineteenth-century inertial frame concept uncritically assumes that two inertial frames agree on whether spatially separated events happen simultaneously. He showed that determining whether two spatially separated events are simultaneous depends on a process of signalling. The velocity of light—implicit in Maxwell's theory and established experimentally by Michelson, Morley, and others—is the same in all reference frames, and Einstein showed that a criterion involving emitted and reflected light signals permits us to identify the time of occurrence of spatially separated events and to derive the Lorentz transformations. This forms the basis of Einstein's special theory of relativity. With Einstein's analysis of simultaneity, the nineteenth-century concept gave way to the 1905 inertial frame concept: An inertial

[45]It is significant that, though Friedman takes inspiration from Poincaré, 'raised to the status of a postulate' means more than simply treating something as a definition. For Einstein, 'raised to the status of a postulate' means treating an empirical principle, e.g, the light postulate, as fundamental and not further explicable; Einstein then uses the light postulate as the basis for a definition or criterion of simultaneity.

frame is not merely one in uniform rectilinear motion but one, furthermore, in which light travels equal distances in equal times in arbitrary directions.

But no sooner was the 1905 inertial frame concept established than Einstein subjected it to a further critical analysis. In 1907, Einstein had an insight that is summarised in the equivalence principle. It is with this principle that the argument for curvature begins.[46]

Before addressing the 1907 insight, however, it is important to note that by 'the equivalence principle,' some will think immediately of the *universality of free fall* that was first established by Galileo: All bodies fall with the same acceleration in the same gravitational field. It may also be stated: The trajectory of a body in a given gravitational field will be independent of its mass and composition. Yet another statement with the same empirical content arises in the framework of Newtonian theory. As is well known, Newtonian theory comprises two different concepts of mass: inertial mass m, the quantity that figures in the second law, that is, the measure of a body's resistance to acceleration; and gravitational mass μ, the quantity that figures in the inverse-square law. It is a well-established experimental fact that the ratio of gravitational mass to inertial mass is the same for all bodies to a high degree of accuracy. And, once we accept that the ratio is a constant, we can choose to use units of measurement that make the two masses for any body equal, so that $\mu/m = 1$. In this way we can ignore the distinction between gravitational mass and inertial mass. This is summarised in what is often called 'the weak equivalence principle': Inertial mass is equivalent to gravitational mass. It is easy to show—though I will not sketch the argument here—that this statement implies that the acceleration of any body due to a gravitational field is independent of its mass and composition.

In Newtonian theory, the proportionality of inertial mass and gravitational mass is a remarkable fact that lacks an explanation. That explanation is found in Einstein's 1907

[46] I intend 'the argument for curvature' as shorthand for 'the arguments for curvature' or, better, 'Einstein's chain of reasoning.' I acknowledge that it may not be possible to formulate the motivation for curvature as a single, coherent argument.

insight of the equivalence principle. The insight is illustrated most clearly with Einstein's 'box.'[47] Suppose you sit in a box from which you cannot look out. You feel a 'gravitational force' towards the floor, just as you would at home. But you have no way of excluding the possibility that the box is part of an accelerating rocket in free space, and that the force you feel is an accelerative force. This also runs the other way: You are inside the box. You feel no gravitational force, just like in free space. But you have no way of excluding the possibility that you are freely falling in a gravitational field. Einstein's insight is that accelerative and gravitational forces must be identical. The insight is summarised in the *equivalence principle*: It is impossible to distinguish locally between a homogeneous gravitational field and a uniformly accelerated frame. This has as a consequence that matter obeys the same laws in a freely falling frame that it would in an inertial frame. Einstein began to recognise in this consequence that inertial motion and freely falling motion are different presentations of the same motion.

But Einstein's argument does not end here: It is crucial that not only matter but light—and moreover, all physical processes—obey the same laws in a freely falling frame that they would in an inertial frame.[48] Einstein's bold extension is motivated by the observation that there are *no* physical phenomena that are independent of gravitation and that could distinguish a box in a homogeneous gravitational field from a box subject to a uniform acceleration. This is also readily illustrated by Einstein's 'box.' Suppose you sit in the box, only this time there is a window. You feel a 'gravitational force' towards the floor, just as you would at home. And, as before, there is no way of excluding the possibility that the box is part of an accelerating rocket in free space, and that the force you feel is an accelerative force. But this time a light ray enters the window. Since light carries energy and energy has mass, the light ray, on entering the box, will not travel across the box horizontally to hit a point opposite its point of entry, but will curve downwards towards the floor—in analogy with a ball thrown horizontally in the gravitational field of the Earth. Assuming that the slight curve of its path were

---

[47]This illustration is found notably in *The Evolution of Physics* (Einstein & Infeld, 1938, pp. 218-222).
[48]This extension of the principle to all physical processes is often referred to as the *universal coupling* of all non-gravitational fields to gravitation.

measurable, a light ray cannot distinguish the box on Earth from the box that is part of the accelerating rocket.

Einstein's insight of 1907, together with this bold extension, led him to recognise freely falling motion and inertial motion as different presentations of the same motion. In this way, the equivalence principle functions as a criterion for identifying two previously distinct concepts of motion.[49]

To return, for a moment, to the proportionality of inertial and gravitational mass in Newtonian theory, *the equivalence principle establishes that homogeneous gravitational forces and accelerative forces are identical*.[50] Since the two concepts of mass figure in the expressions for gravitational force and accelerative force, the principle implies that *inertial and gravitational mass are not merely proportional or equivalent but identical*. In this way the equivalence principle explains the remarkable proportionality of inertial and gravitational mass in Newtonian theory.[51]

Einstein's box illuminates the equivalence principle in both its destructive and constructive aspects. The principle is destructive because it fatally undermines the determinateness of the 1905 inertial frame concept. That is to say, the concept fails to

---

[49]One could object that Newton had already recognised freely falling motion and inertial motion as different presentations of the same motion; one could suggest that Corollary VI to the laws of motion reflects just this. Corollary VI holds that if bodies moving with respect to one another are influenced by uniform accelerative forces along parallel lines they will move with respect to one another in the same way they would if they were not influenced by those forces. In this way, Corollary VI establishes that matter obeys the same laws in a freely falling frame that it would in an inertial frame. But, for Newton, a 'Corollary VI frame' is only an approximation to an inertial frame determined by the laws of motion; it is a good approximation to an inertial frame in the case where the uniform accelerative forces act along lines that are very nearly parallel. Newton had good reason for thinking that the Corollary VI frame should not be identified with an inertial frame. It was Einstein's insight of 1907, and moreover the extension to all non-gravitational forces, that was the crucial interpretive step, namely, recognising freely falling motion and inertial motion as different presentations of the same motion.

[50]Note that the accelerative forces in question here do not include electromagnetic forces or the weak or nuclear forces.

[51]This way of presenting the explanation of the proportionality of inertial and gravitational mass serves to reinforce the importance of prising the universality of free fall from the equivalence principle. Doing so contributes to our understanding of different aspects of the gravitational interaction and to our understanding of the relation between them. But it is essential to note that the so-called equivalence principle is an interpretive extrapolation. The principle that is tested is the universality of free fall.

provide empirical criteria for identifying a unique state of motion as it was supposed to. It is constructive because it motivates a new concept of inertial motion.

With the recognition of a new concept of inertial motion, the question arises: How is this concept to be interpreted? Special relativity presupposes the mathematical framework of an affine space equipped with a Minkowski metric. And, in the special theory, the trajectories of bodies moving inertially as well as those of light rays are interpreted as the straight lines or geodesics with respect to the Minkowski metric while gravitation is a force that pulls bodies off their straight-line trajectories. But Einstein, with the help of Grossmann, saw that Riemann's newly-developed theory of manifolds offered an alternative to such an affine space for interpreting inertial trajectories: Inertial trajectories can be interpreted as the geodesics with respect to a new metric that is determined by the distribution of mass and energy in the universe. Einstein's reinterpretation of free fall is summarised in what is sometimes referred to as the *geodesic principle*: Free massive point-particles traverse time-like geodesics.[52]

With this reinterpretation of inertial trajectories, gravitation is no longer a force causing acceleration, as in Newton's theory, but a manifestation of curvature depending on the mass-energy distribution. This 'geometrisation' of gravitation is at the heart of Einstein's proposal for a new gravitation theory. And, with it, Einstein was faced with the problem of constructing a new theory in which a yet-undetermined quantity representing chrono-geometry is coupled to a yet-undetermined source-term representing the local mass-energy distribution.

The preceding account is a rational reconstruction that avoids various pitfalls and distractions raised by the actual history: from the special theory understood in three-plus-one dimensions *not* four, through Mach's principle and the equivalence principle, the 'rotating disks' and Gauss's theory of non-Euclidean continua, to Riemann's theory and

[52]The geodesic principle is stated in terms of point-particles because it holds only approximately for extended bodies. The geodesic principle for light rays may be stated: Light rays traverse light-like geodesics.

the geodesic principle. However the actual argument falls short, it remains that it was sufficient for motivating a new and purely local definition of a geodesic.

4.2. The equivalence principle and Riemann's theory are not constitutive

With this brief presentation of the argument for curvature in hand, let us return to Friedman's claim that the equivalence principle and Riemann's theory are constitutive.

In my presentation of the argument for curvature, I have shown that the equivalence principle functions as a criterion for identifying two distinct concepts of motion. This identification is the pivotal step that permits the reinterpretation of free fall as geodesic motion. On my analysis therefore—and in contrast with Friedman's—the equivalence principle is not a constitutive principle. Though the principle motivates a new concept of inertial motion, it does not constitute that new concept by expressing a criterion for its application. It is the *geodesic principle* that does that: If a body is freely falling, it is moving on a geodesic; if not, its motion deviates from a geodesic—in a way that a yet-to-be-constructed theory might measure. The geodesic principle forms the basis for treating the relative accelerations of freely falling particles, which can of course be treated in the Newtonian fashion, as a measure of curvature, expressed as geodesic deviation. In this way the geodesic principle replaces the laws of motion as constitutive presuppositions of the concept of inertial motion. The geodesic principle forms the basis for thinking about gravitation as a metrical phenomenon; in other words, for establishing its geometric character. It determines a new framework of investigation, one that makes it possible to pose a question to which Einstein's equations are an answer.

This account is significant for its clarification of the role of the equivalence principle in the conceptual framework of gravitation theory. It also distinguishes the equivalence principle and the geodesic principle as separate elements of that framework. Though the two principles are closely related in Einsteinian gravitation, it is conceivable that future work will reveal that the equivalence principle holds in the face of still more rigorous tests, but that the geodesic principle must be given up—for example, in some new theory of the gravitational interaction.

What of Friedman's claim that Riemann's theory of manifolds is a constitutive presupposition of Einstein's reinterpretation of inertial trajectories as geodesics? Friedman wishes to draw attention to the crucial step of taking spaces of variable curvature to be physical possibilities. It is this step that makes a theory that associates a space-time of variable curvature with the distribution of mass-energy an intellectual possibility. The importance of this step to the construction of the gravitation theory cannot be overstated. But I believe one must distinguish between two things. The first is the transition from the *conceptual framework* of homogeneous spaces—those in which the principle of free mobility is satisfied—to the more general framework of variably-curved spaces in which the former is a special case. The second is the transition to the *mathematical framework* of Riemann's theory of spaces of arbitrarily variable curvature that may be regarded as a realisation of that conceptual framework. While both transitions are prerequisites for the construction of Einsteinian gravitation, it is the transition to the conceptual framework of variably-curved spaces that seems to capture Friedman's point. That is, it is the conceptual framework of variably-curved spaces and *not* Riemann's theory that is constitutive in Friedman's sense. Nor is Riemann's theory constitutive in the narrower sense I have defended. It is, rather, part of the formal background that makes the construction of Einsteinian gravitation possible—in the same sense as the calculus and Euclidean geometry in the case of Newtonian theory. We need some physical principle that expresses criteria for the application of Riemann's theory.

Friedman's inclusion of Euclidean geometry, Riemann's theory, and the calculus in the category of constitutive principles widens that category in the direction of taking everything required for the formulation of a theory to be constitutive. The principles that are truly constitutive are not those that supply the formal background or language necessary for the formulation of a theory or that make particular kinds of applications possible, but those that interpret theoretical concepts by expressing criteria for their application; those same principles control the application of mathematical theories such as Euclidean geometry, affine space, Riemann's theory, and others.

As with my criticism of Friedman's characterisation of the constitutive basis of Newtonian gravitation, this account of the constitutive basis of Einsteinian gravitation in no way undermines Friedman's criticism of Quine's naturalism. I am arguing only for a different account of that basis and a more developed response to Quine. Friedman's account of the structure of physical theories aims to distinguish a theory's constitutive principles from the properly empirical hypotheses whose formulation they permit; and it aims, in this way, to vindicate something close to the analytic-synthetic distinction rejected by Quine. But Friedman's characterisation of both mathematical principles and coordinating principles as constitutive principles neglects that the coordinating principles are answerable to the world while the mathematical principles are not. So, while Friedman is correct to separate constitutive principles from empirical hypotheses, there is a further distinction that his account does not capture.

Let me briefly address Friedman's view that Einstein 'elevated' the equivalence principle to the status of a definition. The idea of such an elevation is based on a misunderstanding of Einstein's 1907 insight that is summarised in the equivalence principle. From my presentation of the argument for curvature, it should be clear that the 1907 insight has nothing to do with an elevation to a definition, but consists in the recognition that inertial motion and freely falling motion are different presentations of the same motion. While the recognition of their identity was the first step in Einstein's argument for a new inertial structure, it seems odd to characterise the principle that brought it about as based on a stipulation ('elevated to a definition'). Provided that one accepts a straightforward fact-convention or fact-definition distinction, the equivalence principle falls clearly on the side of the factual: The universality of free fall is an inductive generalisation from a set of empirical facts, and the equivalence principle is an interpretive extrapolation from the universality of free fall. If any principle were to be elevated, in Friedman's sense, that principle would be the geodesic principle and not the equivalence principle.

4.3. An objection to taking the geodesic principle to be constitutive

There is a possible line of objection to the idea that the geodesic principle is a constitutive principle. It might be pointed out that spinning bodies do not move according to the geodesic principle:

> It has long been recognized that spinning bodies for which tidal gravitational forces act on its elementary pieces deviate from geodesic behaviour. What this fact should clarify, if indeed clarification is needed, is that it is not simply *in the nature* of force-free bodies to move in a fashion consistent with the geodesic principle. (Brown, 2005, p. 141)

But the fact that the geodesic principle is an idealisation—it is strictly satisfied only in the case of zero tidal forces—does not undermine the characterisation of the principle as a constitutive principle. In fact, the idealisation is essential. It is precisely this idealised conception of motion that is the basis for measuring geodesic deviation, which, in Einstein's theory, can be understood in terms of components of rotation, expansion, and shear, given some congruence of geodesics.

It is important to note that an idealised conception of geodesic motion is equally essential to Newtonian theory. The third law of motion asserts that the bodies comprising an isolated system—as well as the particles comprising a single body—will interact with each other so that the forces between them are balanced. In such a state of equilibrium, the centre of mass of the system will follow an approximately geodesic trajectory. The geodesic motion of the centre of mass of an isolated or 'near enough' isolated system is not a precise relativistic notion, but it is crucial to Newtonian reasoning: This state of motion is the basis from which perturbations can be measured.

Newton's method consists in beginning with idealised simple cases and moving to increasingly more complicated ones. In the case of bodies subject to inverse-square centripetal forces, Newton considers in Book I of *Principia*: one-body problems; two-body problems, subject to the third law of motion; and problems of three or more interacting bodies, for which Newton obtains only limited, qualitative results. A distinctive feature of this kind of reasoning is its focus on systematic deviations from Kepler's laws. Smith writes:

> Newton is putting himself in a position to address the complexity of real orbital motion in a sequence of successive approximations, with each approximation an idealized motion and systematic deviations from it providing evidence for the next stage in the sequence. (Smith, 2002, p. 155)

What the work of Smith and others clarifies is that the framework of the laws is the basis for a *perturbative analysis* of planetary systems. That is, the laws are not only a basis for determining the centre of mass of a quasi-isolated system but for reasoning from such a system to a larger system in which the quasi-isolated system is contained and in which systematic deviations from its ideal state of motion can be detected and measured. In both Newtonian and Einsteinian gravitation, therefore, the idealised conception of geodesic motion is the basis for the empirical measurability of the gravitational field. It is the basis for learning about the sources of the gravitational field in the Newtonian picture or for learning about curvature from the relative accelerations of geodesic trajectories in the Einsteinian one.

5. The Kuhnian and Carnapian aspects of Friedman's thesis

In this final section, I wish to consider a further implication of Friedman's view. While Friedman's thesis is primarily motivated as an alternative to Quine's naturalism, it is also a corrective to Kuhn's account of the transition from Newtonian to Einsteinian gravitation. The transition from Newtonian to Einsteinian gravitation is the main example considered by Kuhn in Chapter IX of *The Structure of Scientific Revolutions* (1962 [1970]), and Friedman sees in the logical empiricists' approach to the analysis of physical theories a basis for correcting Kuhn.

In *The Structure of Scientific Revolutions* (1962 [1970]), Kuhn introduced the idea of a scientific 'paradigm,' which he understood not merely as a set of theoretical principles but as an entire world-view consisting of metaphysical views, methodological rules, a conception of what constitutes a legitimate scientific question and what does not, and an understanding of what constitutes a scientific fact. Kuhn called the science pursued within a paradigm 'normal science.' Normal science proceeds without any questioning of basic principles, and consists of puzzle solving, that is, answering questions set by the paradigm with standard methods. Periods of normal science are broken by periods of 'revolutionary science,' which are marked by an accumulation of unsolved puzzles, decreasing confidence in the reigning paradigm, and the appearance of alternative paradigms. Kuhn claimed that science progresses not cumulatively but by a succession of revolutions called 'paradigm shifts.' The main problem posed by this characterisation is this: How can one argue for and commit oneself to a new paradigm if, in periods of revolutionary science, the very criteria of factuality and scientific rationality are being challenged? Kuhn's answer is that the argument for a new paradigm is necessarily circular: 'Each group uses its own paradigm to argue in that paradigm's defense.' (Kuhn, 1962 [1970], p. 94) Paradigm shifts cannot therefore be the result of a rational process; a paradigm shift is ultimately a social or psychological phenomenon. Supposing that one accepts the problem and the response, Kuhn's view can be understood to support relativism, though Kuhn himself did not endorse that consequence.

Friedman's thesis provides an alternative to Kuhn's characterisation of revolutionary theory change. It is distinguished from Kuhn's characterisation in two important respects: its transcendental character, and its replacement of a paradigm shift with a rational process of revision. By its transcendental character, I mean its employment of a method of analysis whose aim is to uncover the principles that interpret theoretical concepts by expressing criteria for their application, and so determine the framework of empirical investigation. It is the revision of *these* principles especially—principles that make possible properly empirical hypotheses, with their associated ontological pictures, methodological rules, puzzles, standards of solution, and modes of community life—that represent revolutionary theory change. Friedman is concerned with the conceptual prerequisites for a theory capable of supporting a tradition of normal science. It would be a mistake therefore to regard the replacement of a set of constitutive principles as an explication of a paradigm shift, even though Kuhn (2000, pp. 104) regarded his own account as 'Kantianism with movable categories.' The replacement of such a set completely replaces Kuhn's idea in an altogether different account of our knowledge and its revision.

It is important to note that, though I intend to criticise the scope of Friedman's characterisation of a constitutive principle as well as his analyses of Newtonian and Einsteinian gravitation, his proposal that we should understand revolutionary theory change as the revision of constitutive principles is not undermined. This aspect of Friedman's account remains even if one accepts my argument that a theory's formal background is not constitutive.

The second respect in which Friedman's proposal differs from Kuhn's concerns the process by which a set of constitutive principles is revised. This part of Friedman's view is subject to the same problems as conventionalism. Friedman has a broadly Carnapian view of theory change and his view inherits something of Carnap's conventionalism. Though Friedman acknowledges Einstein's reinterpretation of free fall, he is more concerned with an external question about the adoption of a new framework than with Einstein's insight within the old framework; but it is this insight that actually motivates the revision. DiSalle (2006b, p. 208) has observed that the question 'do freely falling bodies follow space-time geodesics?' is either an internal question about how geodesics are interpreted in Newtonian theory, in which case it is answered by a mathematical investigation, or an external question about the expediency of adopting a framework in which the trajectories of freely falling bodies are interpreted as geodesics. But, in the context of theory construction, there is no theory in which the trajectories of freely falling bodies are interpreted as geodesics. There is at most the framework of empirical investigation constituted by the geodesic principle—a framework that has yet to lead to the field equations, which, in turn, are a long way from being confirmed. That framework provides us, nonetheless, with a picture of motion, one in which we may ask, for example: What conditions are required for constructing a theory in which free fall trajectories are geodesics? What assumptions must be made about the form of such a theory for Newtonian gravitation to be recoverable in a certain regime? But, with only the external question of whether to adopt Newtonian or Einsteinian gravitation, no such considerations enter into the account of theory change. As DiSalle has put it, 'Carnap's distinction ... does not comprehend the possibility of a conceptual analysis that discovers, within a given framework, the principle on which a radically new framework can be constructed.' (DiSalle, 2006b, p. 208) In the absence of such a possibility, the mechanism of theory change lies in the decision to adopt a framework on the basis of expediency.

Where Carnap's account fails, Friedman's account of the role of distinctly philosophical analysis at a meta-framework level is meant to be a solution. Friedman argues that this distinctly philosophical analysis in periods of fundamental conceptual revolution, in periods when the usual criteria of scientific rationality break down, involves another kind of rationality altogether. This 'communicative rationality' is characterised, roughly speaking, by a process of argument that appeals to patterns of argument acceptable to all participants, with a view to achieving agreement on what the constitutive principles of some domain are. It is opposed to 'instrumental rationality,' which is characterised as an individual process of deliberation in view of achieving some goal. It is the exercise of this form of rationality in both normal and revolutionary science that, for Friedman, typifies Kuhn's failure to find permanent criteria and values across the development of science that enable paradigm shifts to be the result of a rational process. Friedman claims that it is the exercise of communicative rationality that permits agreement on a new framework when framework-dependent criteria of rationality are no longer of service. This is what effects theory change on Friedman's account.

Though Friedman's account of theory change is a significant improvement over Kuhn's, it is still reminiscent of conventionalism. Though it restores the idea that theory change is the result of a rational process and dispenses with mere expediency, the transition from the constitutive basis of Newtonian gravitation to that of Einsteinian gravitation amounts to a conventional choice—a choice achieved by the exercise of communicative rationality. In constrast with Kuhn, Carnap, and Friedman, a better and still more strictly empiricist account of revolutionary theory change is possible. The proper development and defence of this account is beyond the scope of this essay; see DiSalle (2002; 2006a; 2010) and Demopoulos (2010; 2011). But such a development and defence must relocate the role of distinctly philosophical or conceptual analysis: It ought not to be understood as floating above the existing framework and a candidate-framework, which somehow or other have come to be, but as situated in the existing framework, where its objects are those concepts whose interpretations are at issue. To return to the example of the transition from Newtonian to Einsteinian gravitation, the equivalence principle does not merely suggest that the 1905 inertial frame concept may not be the whole story; it undermines the determinateness of the concept definitively and irrevocably. And the consequent reinterpretation of inertial motion as movement along a geodesic that is summarised in the geodesic principle is not a side-effect or by-product of theory change but is itself constitutive of a new framework of investigation. On this understanding, the transition from Newtonian to Einsteinian gravitation is the outcome of a *dialectical* process that begins within the old framework and, through a rational process involving scientific and philosophical considerations, results in a new constitutive principle.

Where Carnap's account cannot comprehend the possibility of an argument for a new framework that has its origin in the old one, this account begins squarely within the old framework. And, by beginning within the old framework, Kuhn's claim that defenders of different paradigms live in different worlds and so cannot argue with each other is undermined.

6. Conclusion

I set out to explicate and evaluate Friedman's thesis. I began by considering Friedman's charactersation of a constitutive principle as well as its antecedents in the work of Reichenbach and Carnap. I proposed that a constitutive principle be characterised as a principle that interprets a theoretical concept by expressing a criterion for its application. And, with this proposal, I argued that the scope of Friedman's characterisation should be narrowed; specifically, that only those principles that have this function should be considered constitutive.

Another task was the evaluation of Friedman's analysis of Einsteinian gravitation. I criticised his claim that the equivalence principle is a constitutive principle. I argued that the equivalence principle is not a constitutive principle but an empirical hypothesis that motivates a new constitutive principle, namely, the geodesic principle. Then I addressed the possible challenge that since free particles follow geodesics only approximately the idea that we should regard the geodesic principle as constitutive is undermined.

The final task was to address the mechanism of theory change in Friedman's view. Though, for Friedman, revolutionary theory change is the result of a process of rational revision and not a paradigm shift, I argued that his idea of a 'change of constitutive principles' is too close to a 'change of conventions.' It is more concerned with the external question of adopting a new framework than with the insight that motivates the revision. I proposed that the role of distinctly philosophical analysis be relocated: It must be situated within the old framework, where the argument for a new constitutive principle begins. In spite of these criticisms, I hope to have shown that Friedman's thesis—at least so far as the methodological analysis of space-time theories is concerned—is eminently defensible. More generally, I aimed to clarify the sense in which Friedman's thesis embraces the transcendental method of analysis without being committed to rescuing Kant's philosophy. Essential to this method of analysis is the recognition that there is a stratification of our knowledge. The idea of a set of constitutive principles stands at some remove from Kant's absolute 'necessities of thought,' but it is concerned nonetheless with the identification of those principles that secure our basic physical knowledge, that make it possible for objects of knowledge to be objects of knowledge. These principles do not have the same status as empirical hypotheses; they are prior to them in that they constitute the framework of empirical investigation, and so make genuine empirical hypotheses possible. This is the aspect of the logical-empiricist approach to the analysis of theories that Friedman seeks to rehabilitate, and that he urges against Quinean and post-Quinean thought.

Looking towards future work, Friedman's thesis is intended to illuminate our analysis of revolutionary theory change not only in space-time physics but in physics and in the other exact sciences. Whether and to what extent this is possible is an open question, as Friedman himself (2001, pp. 117-129) acknowledges. This question is important not only for the further evaluation of Friedman's thesis but, more importantly, for the continuing articulation and evaluation of the idea that Kant's transcendental method is a 'model of fruitful philosophical engagement with the sciences' (Friedman, 1992, p. xii).

References

Aiton, E. (1972). *The Vortex Theory of Planetary Motions*. Macdonald: London.

Brown, H. (2005). *Physical Relativity: Space-time Structure from a Dynamical Perspective.* Oxford: Oxford University Press.

Carnap, R. (1966). *Philosophical Foundations of Physics*. New York: Basic Books.

Carnap, R. (1956). The Methodological Character of Theoretical Concepts. In H. Feigl & M. Scriven (Eds.), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis. Minnesota Studies in the Philosophy of Science, Vol. 1* (pp. 38-76). Minneapolis: University of Minnesota Press.

Carnap, R. (1951). *The Logical Syntax of Language* (A. Smeaton, Trans.). London: Routledge & Kegan Paul. (Original work published in 1934)

Carnap, R. (1939). *Foundations of Logic and Mathematics*. Chicago: University of Chicago Press.

Demopoulos, W. (2013). *Logicism and its Philosophical Legacy*. Cambridge: Cambridge University Press.

Demopoulos, W. (2011). On Extending 'Empiricism, Semantics, and Ontology' to the Realism-Instrumentalism Controversy. *The Journal of Philosophy, 108*, 647-669.

Demopoulos, W. (2010). Effects and propositions. *Foundations of Physics, 40*, 368-389.

DiSalle, R. (2010). Synthesis, the Synthetic A Priori, and the Origins of Modern Space-Time Theory. In M. Domski & M. Dickson (Eds.), *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science* (pp. 523-552). Chicago: Open Court.

DiSalle, R. (2006a). *Understanding Space-Time.* Cambridge: Cambridge University Press.

DiSalle, R. (2006b). Conventionalism and Modern Physics: A Re-assessment. In E. Carson & B. Falkenburg (Eds.), *Intuition and the Axiomatic Method* (pp. 181-212). Dordrecht: Kluwer Academic Publishers.

DiSalle, R. (2002). Reconsidering Kant, Friedman, Logical Positivism, and the Exact Sciences. *Philosophy of Science, 69*, 191-211.

Einstein, A. & L. Infeld. (1938). *The Evolution of Physics*. New York: Simon and Schuster.

Euler, L. (1775). *Lettres à une princesse d'allemagne sur divers sujets de physique et de philosophie*. Londres: chez la Société Typographique. (Original work published in 1768-1772)

Friedman, M. (2011). Einstein and the a priori. Unpublished manuscript.

Friedman, M. (2010a). Synthetic History Reconsidered. In M. Domski & M. Dickson (Eds.), *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science* (pp. 571-813). Chicago: Open Court.

Friedman, M. (2010b). A post-Kuhnian approach to the history and philosophy of science. *The Monist, 93*, 497-517.

Friedman, M. (2006). Carnap and Quine: Twentieth-century echoes of Kant and Hume. *Philosophical Topics, 34*, 35-58.

Friedman, M. (2001). *Dynamics of Reason: The 1999 Kant Lectures of Stanford University.* Stanford: CSLI Publications.

Friedman, M. (1992). *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press.

Kuhn, T. (2000). *The Road since Structure*. Chicago: University of Chicago Press.

Kuhn, T. (1970). *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press. (Original work published in 1962)

Quine, W. V. (1951). Two dogmas of empiricism. *The Philosophical Review, 60*, 20-43.

Reichenbach, H. (1969). *Axiomatization of the Theory of Relativity* (M. Reichenbach, Trans.). Berkeley: University of California Press. (Original work published in 1924)

Reichenbach, H. (1958). *The Philosophy of Space and Time* (M. Reichenbach & J. Freund, Trans.). New York: Dover. (Original work published in 1928)

Reichenbach, H. (1920). *The Theory of Relativity and A Priori Knowledge* (M. Reichenbach, Trans.). Berkeley: University of California Press. (Original work published in 1965)

Smith, G. (2002). The Methodology of the *Principia*. In I. B. Cohen & G. Smith (Eds.), *The Cambridge Companion to Newton* (pp. 138-173). Cambridge: Cambridge University Press.

Wilson, C. (1992). Euler on action-at-a-distance and fundamental equations in continuum mechanics. In P. M. Harman & A. Shapiro (Eds.), *The Investigation of Difficult Things: Essays on Newton and the History of the Exact Sciences* (pp. 399-420). Cambridge: Cambridge University Press.

Chapter 4

On identifying background-structure in classical field theories†

1. Introduction

This essay examines the origin and extension of the concept of background-structure in classical field theories. The extension of the concept, before the recent work of Smolin (2006), Belot (2011), and others, was easily circumscribed. The concept denoted what is characteristic of the space-time structures of Newtonian theory and special relativity.

Newton's laws express criteria of causal interaction. They articulate an account in which the physical quantity *force* is the cause of the acceleration of mass. The content of the laws can be summarised as follows: Given a system of particles in motion, there exists a reference frame and a time-scale relative to which every acceleration is proportional to and in the direction of the force applied, and where every such force belongs to an action-reaction pair.[53] Furthermore, given such a reference frame, forces and masses, accelerations and rotations have the same measured values whether that frame is at rest or in uniform translatory motion. In other words, the laws of motion satisfy the Galilei-Newton relativity principle. The equivalence-class structure determined by the invariance of the laws under the Galilean transformations is the structure of Newtonian space-time.

From a retrospective point of view, the structure of the space-times of the Newtonian and special-relativistic frameworks can be equally well discussed in terms of the mathematical structures that their classes of inertial frames presuppose. In the Newtonian framework, those frames presuppose a global affine structure and separate metrical structures for space and time; in the special-relativistic one, they presuppose global affine and conformal structures and also the metrical structure of space-time. In both frameworks, these structures are fixed independently of the theories of special systems, and thus these structures do not evolve along with the special systems. To use a

---

†A version of this chapter has been published in *Philosophy of Science, 78*(2011): 1070-81. The present version reflects further conversations on this topic with a number of people, especially Erik Curiel, Bill Demopoulos, Robert DiSalle, and Wayne Myrvold.

[53] I owe this formulation to Thomson (1884, p. 387) and Muirhead (1887, pp. 479-480).

common figure, space and time are the ‘stage’ on which the ‘actors,’ namely, the physical fields, move.

One of the great empirical claims of Einsteinian gravitation is that space-time structure is dynamical, and thus something to be discovered empirically. Einsteinian gravitation comprises affine, conformal, and metrical structures. But, in contradistinction to Newtonian theory and special relativity where those structures are necessary presuppositions of the classes of inertial frames, they are fixed only locally, and their variation over any finite region is determined by the distribution of mass-energy. This is not to say that everything in Einsteinian gravitation is dynamical, but, in this way and others, Einsteinian gravitation motivates the revision of the space-time structures of Newtonian theory and special relativity. Space and time cease to be a fixed stage and become actors.

A number of physicists and philosophers of physics, notably Rovelli (2001; 2004), Smolin (2006), and others pursuing loop quantum gravity, have seen in this empirical claim an insight about nature that ought to be preserved in a future theory. In their interpretation of the claim, they have fashioned a new concept that they call ‘background-independence.’ This is the concept to be explicated, but, roughly speaking, to say that a physical theory is background-independent means that physical processes do not unfold against a spatio-temporal framework that is presupposed a priori but determine a dynamical framework in their evolution. This new concept figures in a new heuristic principle that they believe to be fruitful for those pursuing a quantum theory of space, time, and gravitation. Smolin states it as a maxim: ‘Seek to make progress by identifying the background structure in our theories and removing it, replacing it with relations which evolve subject to dynamical law.’ (2006, p. 204). The proper methodological analysis of such a heuristic principle is an outstanding philosophical project, one I hope to pursue in future work.

In this chapter, I take up a prerequisite task. I ask: What is this background-structure that Smolin would have us identify and remove? I propose and evaluate three candidates for background-independence, and I argue that there is something chimaerical about the sought-after concept. My aim, however, is not solely critical and sceptical. I argue that there *is* a proposal for background-independence—one that stems from the work of Trautman, Anderson, and Friedman—that clarifies the particular feature of Einsteinian gravitation that is the basis for nearly all proposals for background-independence.

2. Background-independence and general covariance

There is a sense in which the earliest discussion of background-structure is found in Newton's criticism of Cartesian physics in *De grav*. But let us begin by getting clear on the kind and degree of background-independence exemplified in Einsteinian gravitation.

Einstein took the first steps towards the account of motion characteristic of his gravitation theory in 'On the Electrodynamics of Moving Bodies' (1905 [1952b]). The Newtonian framework uncritically assumes that we have a way of determining whether spatially separated frames agree on which events are simultaneous. In the 1905 paper, Einstein argued that determining the time of occurrence of spatially separated events depends on a process of signalling. The invariance of the velocity of light—implicit in Maxwell's theory and established empirically by Michelson, Morley, and others—provided such a signal, and Einstein argued that a criterion involving emitted and reflected signals permits the derivation of the Lorentz transformations. This is the basis of Einstein's special theory of relativity. One outcome of Einstein's analysis of simultaneity was the replacement of the nineteenth-century inertial frame concept with the 1905 inertial frame concept: An inertial frame is not merely one in uniform rectilinear motion but also one in which light travels equal distances in equal times in arbitrary directions.

With the special theory of relativity, it was necessary to find a new theory of gravitation that would overcome the contradiction between the invariance of the velocity of light and the instantaneous action at a distance postulated by Newtonian gravitation. In 1907, Einstein had an insight that has come to be called the *equivalence principle*. This is the hypothesis that it is impossible to distinguish locally between a homogeneous gravitational field and a uniformly accelerated frame. The equivalence principle has as a consequence that matter obeys the same laws in a freely falling frame that it would in an inertial frame. In this consequence, Einstein began to recognise that freely falling motion and inertial motion are different presentations of the same motion.

But there is a further step in Einstein's argument: Einstein argued from the hypothesis that all bodies fall with the same acceleration in the same gravitational field to the stronger hypothesis that not only matter but light—and moreover, all physical processes—obey the same laws in a freely falling frame that they would in an inertial one. Without this extension, some phenomena, electromagnetic phenomena, e.g., would be a basis for measuring the acceleration of a freely falling particle relative to electromagnetically accelerated trajectories, namely, trajectories not determined by gravitation. This would be no different from our ability to measure the acceleration of an electron in an electromagnetic field relative to the inertial trajectory of a particle that is not affected by that field. This extension reflects what Will (1993, p. 68) has called the 'universal coupling' of all non-gravitational fields to the gravitational field. If any phenomena failed to couple to gravitation in this way, they would indicate the existence of a 'background-structure' that is distinguishable from the gravitational field. As Will has put it, universal coupling allows us to 'discuss the metric g as a property of space-time itself rather than as a field over space-time' (Will, 1993, p. 68).

With his insight of 1907 and the crucial extension to all physical processes, Einstein recognised that freely falling motion and inertial motion are different presentations of the same motion. In this respect, the equivalence principle functions as a criterion for identifying inertial frames and freely falling ones. The identification of classical inertial frames and freely falling ones fatally undermines the determinateness of the 1905 inertial frame concept. That is, the concept fails to provide empirical criteria for identifying a unique state of motion as it was supposed to. With this identification, the fundamental distinction between inertial and non-inertial frames was collapsed, and the relevant distinction became one between systems in free-fall and systems in non-free-fall motion.

With the 1907 inertial frame concept, Einstein was faced with the question, how is the concept to be interpreted? Einstein's chain of reasoning is the subject of debate, but there is a rational reconstruction of that reasoning that highlights the essential steps. Special relativity presupposes the mathematical structure of an affine space equipped with a Minkowski metric. In the special theory, the trajectories of bodies moving inertially and also those of light rays are interpreted as the straight lines or geodesics with respect to the Minkowski metric while gravitation is a force that pulls bodies off their straight-line trajectories. But Einstein, with much help from Grossmann, saw that Riemann's newly-developed theory of manifolds offered an alternative to such an affine space for interpreting inertial trajectories: The inertial trajectories of freely falling particles can be interpreted as the geodesics with respect to a new metric that is determined by the distribution of mass and energy in the universe. This reinterpretation of free fall is summarised in what has been called the *geodesic principle*: Free massive point-particles traverse time-like geodesics.[54] The geodesic principle interprets the 1907 inertial frame concept by expressing a criterion for its application. It provides a framework of investigation in which one can begin to think about how to construct a theory where gravitation is represented as a manifestation of the curvature of space-time structure that is determined by the distribution of mass and energy. In this framework of investigation, Einstein realised that there is no way of smoothly laying down a global coordinate system and that the laws of his gravitation theory required a coordinate-independent expression. He referred to that requirement as the *principle of general covariance*.

With the geodesic principle and the requirement of general covariance with which Einstein connected it, no longer was there an equivalence class of preferred coordinate systems determined by the laws, and no longer were there non-dynamical affine,

[54]The geodesic principle is stated in terms of point-particles because it holds only approximately for extended bodies. The geodesic principle for light rays may be stated: Light rays traverse light-like geodesics.

conformal, and metrical structures. Though Einstein did not use the term 'background-independence,' he certainly appealed to the notion in his own characterisations of his gravitation theory. I will call that notion

> *Proposal 1.* A theory is background-independent just in case it satisfies the requirement of general covariance.

But no sooner do we have this proposal in hand than we must respond to an objection: General covariance was trivialised nearly as soon as it was presented. In 'The Foundation of the General Theory of Relativity' (1916 [1952a]), Einstein gave an argument for general covariance that, following Stachel (1980), we now know as 'the point-coincidence argument.' The locution 'point-coincidence' refers to the view that all physical observations consist in the determination of purely topological relations (coincidences) between objects of spatiotemporal perception. The argument runs as follows: (P1) All evidence for or against a physical theory rests on immediately verifiable facts. (P2) Immediately verifiable facts are exhausted by point-coincidences. (C) Thus, physical observations are reducible to point-coincidences. On this argument, any mapping that preserves point-coincidences preserves a theory's physical content, and thus no coordinate system is privileged.

Kretschmann (1917) brought to light an important physical implication of the point-coincidence argument that he took to trivialise general covariance. He thought that, if indeed a theory's physical content is exhausted by point-coincidences, the equations of any theory can be made generally covariant without a modification of that content. Kretschmann's challenge was taken seriously in the 1960s by Trautman, Anderson, Wheeler, Fock, and others, who learnt to distinguish the requirement of general covariance from the symmetries that equations of motion formulated in the Einsteinian framework admit. Henceforth, we will be discussing those symmetries and not the requirement of general covariance as understood by Einstein.

3. The Anderson-Friedman programme

Anderson (1967) challenged the view that general covariance is the characteristic feature of Einsteinian gravitation, pointing out, as Kretschmann did, that any theory can be given a generally covariant formulation. He claimed that the characteristic feature of Einsteinian gravitation is its lack of an 'absolute object.' Anderson's proposal was taken up by Friedman (1983), who sought to give it a more perspicuous formulation, and the following definitions are Friedman's. To state the proposal properly, I will give an abstract sketch of a classical field theory. I will do so only in meanest outline and in a familiar notation. See Pitts (2006) for a technically and historically careful treatment of the Anderson-Friedman programme and the differences between Anderson's and Friedman's definitions.

Let me represent the *space-time* of a classical field theory T as an ordered n-tuple of the form $(M, O_1, ..., O_n)$, where M is a smooth manifold and $O_1, ..., O_n$ are *geometric objects* on M. Defining geometric objects is a non-trivial task, but, in general, the objects in question are tensors, tensor fields, and also metric-compatible connections. The *dynamical laws* of T will be built up out of these geometric objects. These laws have the form $f(O_1, ..., O_n) = 0$.

Let me turn now to the notion of an automorphic *mapping* of geometric objects on the manifold. If $(M, \phi_1, ..., \phi_n)$ and $(M, \theta_1, ..., \theta_n)$ are both models for T, then for every point p of M there is a mapping d of a neighbourhood A of p onto a neighbourhood B of p such that $\phi_i = d\ \theta_i$ on $A \cap B$. If that mapping is infinitely differentiable, one-to-one, onto, and has an infinitely differentiable inverse, then the mapping, denoted d, is called a *diffeomorphism*. The arbitrary diffeomorphisms d form a group, often denoted *diff(M)* as a reminder that they are automorphisms of M. Elements of *diff(M)* act on the geometric objects of the theory in question.

With this framework in hand, let me return to the Anderson-Friedman proposal for characterising Einsteinian gravitation and other classical field theories. A geometric object O_i is an *absolute object* of T just in case for any two T-models $(M, \phi_1, ..., \phi_n)$ and

(M, θ_1, ..., θ_n) ϕ_i and θ_i are invariant under *diff(M).* A geometric object that does not satisfy this definition is a *dynamical object*.

Anderson's distinction between absolute and dynamical objects is the basis of his definition of a theory's *symmetry group*; namely, the largest subgroup of *diff(M)* that leaves invariant the theory's absolute objects. It is noteworthy that, though Anderson defines a theory's symmetry group in terms of that theory's antecedently defined absolute objects, on an alternative understanding, the lack of absolute objects would be expressed by the lack of non-trivial symmetries.

This definition is significant because it meets Kretschmann's challenge: Theories may be reformulated so that their geometric objects are invariant under the actions of subgroups of *diff(M)* like the Poincaré group or so that they are invariant under *diff(M)* itself, even though, in their standard formulations, they would be invariant only under more limited mapping groups. It is precisely the further requirement expressed in the above definition that is supposed to distinguish a theory's symmetry group from its mapping or covariance group. That requirement distinguishes Einsteinian gravitation, which Anderson claimed lacks an absolute object, from previous theories.

4. Background-structure represented by geometric objects and beyond

I have presented the definition of an absolute object not only to move beyond the trivialisation of general covariance but because Anderson took the presence of absolute objects in a theory's equations to imply that theory's commitment to a certain form of 'background-structure,' though he himself did not use that term. Thus, the Anderson-Friedman definition provides us with another strategy for identifying background-independence, which I will call

> *Proposal 2.* A theory is background-independent just in case it has no absolute objects.

With this proposal, the notion of background is entirely determined by the geometric objects on a manifold.[55]

As with proposal 1, no sooner do we have this proposal in hand than we must respond to a line of objection, namely, that the Anderson-Friedman distinction between absolute and dynamical objects cannot capture the intended and essentially physical distinction. Geroch (reported in Friedman, 1983, p. 59, n. 9) pointed out that even in Einsteinian gravitation one might draw up a scenario in which geometric objects like nowhere-vanishing vector fields and symplectic forms count as absolute objects. He made his point with the following example. Suppose we have a cosmological model in which there is omnipresent dust, all particles of which are at rest in some Lorentz frame. Pressure-free dust has the stress-energy tensor $T^{ab} = \rho U^a U^b$, where the density of the dust particles ρ is defined as the number of particles per unit volume in the unique inertial frame in which the particles are at rest and U^a is the four-velocity. In such a universe, the four-velocity would be nowhere-vanishing, and would count as an absolute object on Friedman's definition. That is, there would be a background reference frame in the imaginary model, the rest frame of the dust. Torretti (1984, p. 285) offered another counterexample to the Anderson-Friedman distinction. He formulated a theory of modified Newtonian mechanics in which each model has a space of constant non-positive curvature, but different models have different values of curvature. He pointed out that such curvature is undeniably a kind of background-structure, yet escapes the Anderson-Friedman definition of absoluteness. Pitts (2006) presents and challenges these and other counterexamples, and he offers a defence of the Anderson-Friedman programme. But he concedes that Einsteinian gravitation may have an absolute object, namely, the scalar

[55]There is a discussion that I would like to acknowledge, if only briefly. Though Anderson did not introduce absolute and dynamical objects with reference to action principles, he certainly regarded dynamical objects as variational, while absolute objects are not (1967, 88-89). Some (e.g., Hiskes, 1984) have seen in this another way of drawing the absolute-dynamical distinction: No object that is varied in a theory's action principle should be considered absolute. But others (e.g., Rosen, 1966; Sorkin, 2002) have argued that a flat metric can be derived from an action principle by introducing geometric objects that vary in the required way. It is significant that Anderson himself (1967, p. 83) headed off this line of objection by proscribing what he called 'irrelevant variables.' Anderson was concerned with the essentially physical distinction between space-time structure in Newtonian theory and special relativity, on the one hand, and Einsteinian gravitation, on the other. For him, that distinction was never a merely formal one, and he was at pains to defend it from those who would undermine it with formal 'tricks.'

density obtained by reducing the metric into a conformal metric density and a scalar density.[56]

Some may consider these counterexamples to be reason enough for giving up proposal 2. But Trautman (1966; 1973) hints at another way of thinking about physical theories, one that Pitts does not consider in his defence.[57] This work suggests that the space-times of Newtonian theory and special relativity are *characterised* by absolute objects; the space-time of Einsteinian gravitation is *characterised* by dynamical ones.[58] That is to say, theories of special systems formulated in the Newtonian or special-relativistic frameworks presuppose geometric objects that determine a fixed metric affine geometry; those of systems formulated in the framework of Einsteinian gravitation depend on geometric objects that determine a dynamical one. On this interpretation, there is no suggestion that Einsteinian gravitation lacks an absolute object. The distinction between Anderson's and Friedman's accounts, on the one hand, and Trautman's, on the other, is not merely verbal. The claim that the metric affine geometry of Einsteinian gravitation is characterised by dynamical objects is importantly different from the claim that Einsteinian gravitation has no absolute objects. In this way, the line of objection motivated by the counterexamples and a debate over the viability of the Anderson-Friedman programme is better avoided. This way of characterising physical theories can also be used to motivate another proposal for background-independence; we might call it proposal 2a: A theory is background-independent just in case its metric affine geometry is characterised by dynamical objects. This proposal helps preserve something of the intended and essentially physical distinction that motivated the distinction between absolute and dynamical objects.

Nonetheless, there is a line of objection that undermines both proposals 2 and 2a in a different way. These proposals commit us to the view that whether a theory is

---

[56]See Pitts (2006, pp. 366-367) for details.

[57]In fact, the notions of absolute and dynamical objects are due to Trautman. But it was Anderson and Friedman who gave them a perspicuous formulation. For this reason, Anderson and Friedman are more readily associated with them than Trautman.

[58]Passages supportive of this reading can be found in Trautman (1973), though the claim that Einsteinian gravitation has no absolute object can also be found in Trautman (1966). In any case, I will attribute this reading to Trautman.

background-independent depends on its geometric objects. But Belot (2011, pp. 12-20) has recently pointed out that the concept of background-independence admits of degrees. He considers, among other examples, the vacuum solutions to Einstein's field equations that give rise to de Sitter, anti-de Sitter, and Minkowski space-times. These and other solutions have the asymptotic behaviour of one of the spaces of constant curvature.

To take another family of examples in Einsteinian gravitation, suppose one attaches a boundary to a four-dimensional manifold.[59] Suppose, further, that one builds into the kinematically possible configurations of a theory's geometric objects not only such requirements as smoothness and global hyperbolicity but also the requirement that space-time is approximately Minkowskian as one approaches the boundary.[60] Such a theory will have no geometric objects that determine a background, but such a theory will admit *diff(M)* only locally, not generally; at the boundary, the theory will admit only a subgroup of *diff(M)*. The theory will lie between paradigmatically background-dependent theories in which geometric objects propagate in Minkowski space-time and paradigmatically background-independent theories such as spatially compact Einsteinian gravitation. So, even though the theory has no geometric objects that determine a background-structure, the boundary conditions ensure that any solution has the structure of a Minkowskian background at spatial infinity.

With these sorts of situations in mind, Belot proposes an elegant scheme for fixing the extension of background-structure.[61] No longer is background-independence an all or nothing affair; theories are shown to have degrees of background-independence. To make precise various degrees of background-(in)dependence, Belot introduces a distinction between a theory's geometrical and physical degrees of freedom. The *geometrical degrees of freedom* are represented by the geometric objects, figuring in the dynamical laws of a theory, that parametrise the equivalence classes of space-time geometries. The *physical degrees of freedom* are represented by the geometric objects that parametrise the quotient-space obtained by identifying gauge-equivalent solutions. A

[59] For details on attaching various kinds of boundaries, see, e.g., Hawking and Ellis (1973).
[60] I owe this family of examples to Belot (2011).
[61] The following is only a sketch of Belot's proposal; see Belot (2011) for details.

theory is then said to be *fully background-dependent* just in case it has no geometrical degrees of freedom, and *fully background-independent* just in case its geometrical and physical degrees of freedom match. Of greater moment, however, is the possibility of characterising theories of ambiguous background-structure. A theory is said to be *nearly background-dependent* if it has only finitely many geometrical degrees of freedom and *nearly background-independent* if it has a finite number of non-geometrical degrees of freedom. In this way, proposals 2 and 2a are recovered and situated in a larger space of possibilities in which their uniqueness is undermined.[62]

Belot's proposal is a significant contribution. But, as much as it clarifies the idea that there are various degrees of background-independence, it also sharpens the ambiguity of the concept. Provided that a theory has no absolute objects, does background-independence require (i) that a theory *presuppose* nothing about global structure or (ii) that a theory *preclude* the possibility of such structure? For Einsteinian gravitation could be said to satisfy neither (i) nor (ii) since the theory holds that geometry is everywhere locally Lorentzian, making Belot's proposal trivially true, or Einsteinian gravitation could be said to satisfy only (i) in that geometry is dependent on the distribution of mass and energy. Though I leave aside the question of the methodological status of a principle like Smolin's for future work, this particular ambiguity already suggests a reason to avoid asserting a meta-principle about eliminating background-structure.

Belot's proposal also provides an opportunity to comment on the distinction between local and global structure in Einsteinian gravitation. Einsteinian gravitation departs from Newtonian theory and special relativity in that it places weaker a priori restrictions on global structure; to put the point in Carnapian terms, global geometry is relegated to the *P*-rules of the framework. But that departure, though radical, does not stem from a philosophical or methodological motivation to construct a theory with that

[62]Belot's proposal represents a significant advance over the work of Trautman, Anderson, and Friedman, but it is noteworthy that the idea that geometric objects parametrise a theory's degrees of freedom is already there in Trautman (1966, p. 322).

characteristic but from the fact that the equivalence principle motivates a purely local definition of a geodesic.

5. Further beyond geometric objects

To this point, I have only considered some of the strongest mathematical structures that may be imposed on a manifold; namely, metrics and other geometric objects both absolute and dynamical. I have considered certain solutions to the field equations and also the imposition of asymptotic boundary conditions from which background-structure may arise. In contrast, Trautman (1973), Thorne, Lee, and Lightman (1973), Smolin (2006), and others have pointed out that one may count dimension, topological and differential structure, temporal orientation, and even the metric signature as background-structures, though they leave it as an open question whether these lower levels of background-structure are essential to all physical theories or whether they may be replaced by a future theory. In this vein, one might ask, why use the real numbers as opposed to some other field? And, taking this still further, one might well ask whether all the mathematical structures a theory 'quantifies over' are to be considered background-structures. Though I take the suggestion of Trautman and others seriously, it reinforces that there is something chimaerical about background-independence: No sooner have we cut off one head than two more spring up to take its place; no sooner do we seem to be getting a hold of the concept when it slips away again. In any case, it is noteworthy that Einsteinian gravitation presupposes these lower-level features, yet allows for scenarios in which certain of these features are violated by (e.g.) singularities.

How, then, are we to fix the extension of background-independence so as to include those kinds of background-structures that escape a proposal such as Belot's? There is an intuition that seems to underlie the views of Smolin (2006), Giulini (2007), Belot (2011), and others. And, though I do not do full justice to their views, I will summarise it in what I call

> *Proposal 3.* A theory is background-independent just in case it has no fixed 'stage' that shapes the evolution of the fields without itself being shaped by them.

This proposal is very nearly the so-called action-reaction principle: For something to be physical it cannot act without being acted upon. That principle, Einstein claimed, is satisfied by his gravitation theory and not by Newtonian theory or special relativity. And the idea certainly lies behind Anderson's definition of an absolute object. I will not address here Einstein's view that space-time should not act without being acted upon. Nor will I address the bearing of the action-reaction principle on discussions of background-independence. But it is important to note that the action-reaction principle seems to loom behind nearly all proposals for fixing the extension of background-independence.

A metaphor by Novalis—'Theories are nets: only he who casts will catch'—is particularly apt for the evaluation of proposal 3. The main objection to proposal 3 is that it is a step too far; it is a catch-all for virtually any kind of mathematics used in the formulation of a theory. All of the above proposals are subsumed, but room is left for other conceptions of background-structure. It may be that some still finer-grained classification is possible, but I will not attempt that here. I only want to point out that, by catching everything, proposal 3 blurs even the line between the *language* required for saying anything at all and *interpreted mathematical theories*. Where that line is drawn varies from theory to theory, and it is not drawn a priori or by some philosophical or methodological demand such as (e.g.) the demand that a theory be background-independent. Rather, it is a set of empirical criteria—the laws of motion, a criterion for identifying time of occurrence, the geodesic principle—that controls the application of some or another body of mathematical theory. In Einsteinian gravitation, for instance, one needs an empirical reason to consider certain solutions to the field equations as physical possibilities or to impose asymptotic boundary conditions, and, in view of that, one might not want to formulate the theory or a meta-theoretical principle about the theory so that certain solutions or the imposition of boundary conditions is precluded a priori. At the very least, an important strength of proposals 2 and 2a over proposal 3—or any proposal motivated by the action-reaction principle or something like it—is that it does not dissolve the important differentiation of background-structures into a ‘night in which all cows are black.’

If proposal 3 is a step too far, what, if anything, remains to be said about proposals 2 and 2a? I have presented the case against these proposals: I have charged them with failing to account for kinds of background-structures that are not determined by the geometric objects on a manifold. To be sure, the analysis of the concept of background-structure cannot end with proposals 2 and 2a. But the net cast by Trautman, Anderson, and Friedman was a good one. It is a virtue of proposal 2a that it illuminates a central feature of Einsteinian gravitation: The Einsteinian framework does not presuppose certain global structures, though it does not preclude them. It does not preclude, for example, that we might want to study bounded systems like stars, and so investigate space-times that are asymptotically flat. That one can formulate and study scenarios such as those identified by Geroch and Belot does not diminish that proposal’s isolation of the difference between theories formulated in a Newtonian or special-relativistic framework, on the one hand, and certain theories formulated in the framework of Einsteinian gravitation, on the other. In this way, that proposal sharpens—and does not blur—the feature of Einsteinian gravitation that is the basis for nearly all proposals for background-independence.

With a clearer understanding of the empirical motivation for a dynamical geometry, we might attempt to express Smolin’s methodological principle more reasonably. We might formulate it: ‘Find out whether there are background-structures that cannot be empirically motivated and eliminate them.’ But this principle, too, reflects no philosophical insight peculiar to Einsteinian gravitation. In the absence of some particular empirical motivation for applying or eliminating a given mathematical structure, it reflects only the standard empiricist’s application of Ockham’s razor—and it could apply not only to background-structures, understood in terms of absolute objects, but also to dynamical objects if they have no empirical motivation. By criticising Smolin’s principle, I do not mean to suggest that no meta-principles play or have played a heuristic role in theory construction. But what is revealed in Einstein’s own construction of his theories, and in his provision of empirical criteria that articulate theoretical concepts, is that *methodological analysis* promises a clearer understanding than *meta-principles* of what constraints are imposed by our present understanding of gravitation on future theories.

6. Conclusions and a further consideration

With each of my proposals, I have tried to identify a genuine candidate for background-independence. There is a sense in which the requirement of general covariance is a candidate for background-independence. But Kretschmann showed that theories that are not generally covariant in their standard formulations may be reformulated, a point that the Anderson-Friedman programme masterfully addressed. I argued next that there is an important sense in which a theory that has no absolute objects—or whose metric affine geometry is characterised by a dynamical object—is a candidate for background-independence. But the challenges to proposals 2 and 2a, from several directions, seemed to suggest that this strategy could not capture important conceptions of background-structure. This motivated proposal 3. That proposal captures too much, and I suggested that one might want to distinguish between different kinds of background-structures; namely, those arising from certain solutions to Einstein's field equations, from the imposition of boundary conditions, and from lower levels of background-structure. In this regard, Belot's proposal is significant for its clarification of the sense in which even a theory whose metric affine geometry is characterised by a dynamical object can still have various degrees of background-independence. But Belot's proposal also draws attention to the more basic question of what is demanded by the concept of background-independence. I have suggested that, though Belot's analysis provides an important explication, it says nothing about whether the concept requires that a theory *presuppose* nothing about global structure or that a theory *preclude* the possibility of such structure. The empirical interest of studying certain isolated systems in Einsteinian gravitation, e.g., would appear to be a good reason not to preclude the possibility of such structure a priori. We are left, then, with only the weaker demand that a theory presuppose nothing about global structure—at least so far as we set aside questions about lower-level background-structures like topological and differential structures, the metric signature, and others. This is the sense that the programme of Trautman, Anderson, and Friedman succeeds in capturing. So, though there is something chimaerical about background-independence, there is also a sense in which the feature of Einsteinian gravitation that motivated all of the proposals is aptly captured by that programme.

There is a further consideration with which I would like to conclude. I set out by recalling the fundamental insight into the nature of the gravitational interaction that is summarised in the equivalence principle. I recalled that the equivalence principle motivated a new inertial frame concept, and that the geodesic principle expresses a criterion for the application of that concept.

The equivalence principle that I have discussed has been called 'Einstein's equivalence principle,' which is itself an interpretive extrapolation of the universality of free fall. Einsteinian gravitation is not the only theory that satisfies this equivalence principle; Newtonian theory satisfies it too. Anderson (1967), Ehlers (1973), and others have shown that Einsteinian gravitation also satisfies another principle. This has been called 'the principle of minimal coupling,' according to which no terms of the special-relativistic equations of motion contain the Riemann curvature tensor. In this way, minimal coupling ensures that special relativity is a local approximation so long as tidal gravitational effects can be ignored. Not only does Einsteinian gravitation satisfy this stronger demand; it is essential for ensuring the local validity of special relativity. The conjunction of Einstein's equivalence principle and minimal coupling amounts to what has been called 'the strong equivalence principle.'[63] The strong principle implies that no more than the dynamical metric g_{ab} is needed to account for gravitation.

This stronger principle bears directly on the question of background-independence because a modification of Einsteinian gravitation like the Brans-Dicke theory, which

[63]Note that the strong equivalence principle is more commonly presented as the conjuction of the principle of the universality of free fall, which grounds Einstein's equivalence principle and is in this sense more fundamental, and the principle of minimal coupling. (See, for example, Anderson (1967), Ehlers (1973), and Will (1993).) I have discussed Einstein's equivalence principle throughout because its role in the argument for the 1907 inertial frame concept is more immediate.

comprises absolute and dynamical objects, fails to satisfy it. The exclusion of the Brans-Dicke theory and others by the strong equivalence principle serves to isolate, in yet another way, what is distinctive about Einsteinian gravitation. Here the strong equivalence principle is playing the same role as proposal 2 in excluding those theories that comprise absolute objects. In this way, the strong equivalence principle further clarifies the sense in which Einsteinian gravitation is background-independent, whether or not the concept is of any service as a heuristic.

References

Anderson, J. L. (1967). *Principles of Relativity Physics.* New York: Academic Press.

Belot, G. (2011). Background-independence. *General Relativity and Gravitation, 43*, 2865-84.

DiSalle, R. (2006). *Understanding Space-Time*. Cambridge: Cambridge University Press.

DiSalle, R. (2002). Reconsidering Ernst Mach on Space, Time, and Motion. In D. Malament (Ed.), *Reading Natural Philosophy*, pp. 167-191. Chicago: Open Court.

Ehlers, J. (1973). Survey of General Relativity Theory. In W. Israel (Ed.), *Relativity, Astrophysics, and Cosmology* (pp. 1-125). Dordrecht: D. Riedel.

Einstein, A. (1952a). The Foundation of the General Theory of Relativity. In H. A. Lorentz, A. Einstein, H. Minkowski, & H. Weyl, *The Principle of Relativity* (pp. 109-164). New York: Dover. (Original work published in 1916)

Einstein, A. (1952b). On the Electrodynamics of Moving Bodies. In H. A. Lorentz, A. Einstein, H. Minkowski, & H. Weyl, *The Principle of Relativity* (pp. 35-65). New York: Dover. (Original work published in 1905)

Friedman, M. (1983). *Foundations of Space-Time Theories*. Princeton: Princeton University Press.

Giulini, D. (2007). Some remarks on the notions of general covariance and background-independence. In E. Seiler & I.-O. Stamatescu (Eds.), *Approaches to Fundamental Physics: An Assessment of Current Theoretical Ideas*, pp. 105-120. Berlin: Springer-Verlag.

Hawking, S. & G. Ellis. (1973). *The Large Scale Structure of Space-time*. Cambridge: Cambridge University Press.

Hiskes, A. L. D. (1984). Space-time theories and symmetry groups. *Foundations of Physics, 14*, 307-332.

Kretschmann, E. (1917). Ueber die physikalischen Sinn der Relativitätspostulaten. *Annalen der Physik, 53*, 575-614.

Pitts, B. (2006). Absolute Objects and Counterexamples. *Studies in History and Philosophy of Modern Physics, 37*, 347-371.

Rosen, N. (1966). Flat space and variational principle. In B. Hoffmann (Ed.), *Perspectives in geometry and relativity: Essays in honor of Václav Hlavaty*. Bloomington: Indiana University.

Rovelli, C. (2004). *Quantum Gravity*. Cambridge: Cambridge University Press.

Rovelli, C. (2001). Quantum Spacetime: What Do We Know? In C. Callender & N. Huggett (Eds.), *Physics Meets Philosophy at the Planck Scale*. Cambridge: Cambridge University Press.

Smolin, L. (2006). The Case for Background Independence. In D. Rickles, S. French, & J. Saatsi (Eds.), *The Structural Foundations of Quantum Gravity*. Oxford: Oxford University Press.

Sorkin, R. (2002). An Example Relevant to the Kretschmann-Einstein Debate. *Modern Physics Letters A, 17*, 695-700.

Stachel, J. (1980). Einstein's Search for General Covariance. In D. Howard & J. Stachel (Eds.), *Einstein and the History of General Relativity (Einstein Studies, Volume 1)*, pp. 63-100. Boston: Birkhäuser.

Thorne, K. S., Lee, D. L., & Lightman, A. P. (1973). Foundations for a theory of gravitation theories. *Physical Review D, 7*, 3563-3578.

Torretti, R. (1984). Review: Space-Time Physics and the Philosophy of Science. *British Journal for the Philosophy of Science, 35*, 280-292.

Trautman, A. (1973). Theory of Gravitation. In J. Mehra (Ed.), *The Physicist's Conception of Nature* (pp. 179-198). Dordrecht: D. Reidel Publishing.

Trautman, A. (1966). The General Theory of Relativity. *Uspekhi Fizicheskikh Nauk, 89*, 3-37.

Will, C. (1993). *Theory and Experiment in Gravitational Physics*. Cambridge: Cambridge University Press.

Chapter 5

Conclusion

In each of the essays, I have addressed the proper methodological analysis of inertial frames in the conceptual framework of physics. By way of conclusion, I would like to consider the notion of apriority that figures, in one way or another, in each of the essays.

5.1. The a priori and the foundations of space-time theories

There are two senses of 'a priori' that arise in the essays. The first sense is found in 'There is no conspiracy' and 'On identifying background-structure.' This sense concerns the legitimacy of accepting a physical theory that asserts the possibility of global space-time structures.

This line of enquiry looks both backwards and forwards. Looking backwards, Brown claims that there is something questionable about Newtonian theory and special relativity because they assert the possibility of global inertial frames, along with the geometric objects required for their representation. This view is driven by arguments that are not exclusively or even primarily empirical. Looking forwards, Smolin and others are concerned with the questions: Can there be any a priori demand on a future theory? Is eliminating 'background-structure,' however this term is understood, a fruitful heuristic in the pursuit of a quantum theory of space, time, and gravitation?

In both of these essays, I argued that the offending non-dynamical background-structures—those characteristic of Newtonian theory and special relativity and supposedly explained away by Einsteinian gravitation—are *not* the result of an a priori postulation. To be sure, there are geometric objects that are invariant in all models of Newtonian theory and special relativity. But the criteria that control the application of these geometric objects are derived from experience. In the absence of an empirical criterion for identifying classical inertial frames and freely falling ones, therefore, the alleged illegitimacy of the global inertial frames of Newtonian theory and special relativity amounts to no more than the expression of a meta-theoretical or metaphysical preference for a certain kind of theory.

The main motivation of the conspiracy of inertia and of the proposals for background-independence appears to stem from a certain approach to the analysis of physical theories. This may consist of taking certain metaphysical or meta-theoretical principles to be satisfied by Einsteinian gravitation. It may consist of taking certain features of Einsteinian gravitation and holding them to be insights about nature that future theories must satisfy. In either case, it amounts to elevating certain ideas to the status of standards according to which both older theories—Newtonian theory and special relativity—and new research programmes—approaches to quantum gravity—are evaluated. At first glance, the tremendous empirical success of Einsteinian gravitation could appear to justify such an elevation. That elevation could be regarded as fulfilling the empiricist demand that experience should make a rational contribution to our knowledge of the world.[64] But, whatever the precise motivation of the conspiracy allegation and of the proposals for background-independence, this approach to the analysis of theories finds defects in the global inertial frames of Newtonian theory and special relativity where in fact there are none. This approach, while seemingly motivated by a commitment to empiricism, is driven by a set of metaphysical and methodological intuitions that go beyond empiricism's basic demand.

The second sense of 'a priori' arises in 'Friedman's Thesis.' The characterisation of a constitutive principle I have defended stands at some remove from Kant's. Their relation is one of analogy: Kant's constitutive principles apply the categories of the understanding to possible experience; constitutive principles, in the sense I have defended, interpret theoretical concepts by expressing criteria for their application. So, though these principles may be called 'constitutive principles' in analogy with Kant's principles, the principles in question are not a priori at all. They are prior only to the properly empirical hypotheses that they make possible. Furthermore, constitutive

---

[64] I owe this characterisation of the empiricist thesis to Gupta (2006).

principles must be distanced from principles that are true in virtue of their meanings, in virtue of convention or stipulation.

The notion of a constitutive principle also figures in Kant's analysis of Newtonian physics. Kant saw in Newton's theory not only a revolutionary scientific discovery but a revolutionary philosophical advance. He saw in Newton's theory a basis for criticising the preceding Leibnizian tradition in which notions of space, time, force, motion, substance, and causality are applied to the 'intelligible' world of monads. He held that these metaphysical notions have no content at all except through their 'sensible' counterparts, and he took the laws of motion to provide the only framework in which the sensible notions can be extended to the universe at large.

While Kant was mistaken in thinking that Newtonian physics is the unique such framework, the idea that certain theories provide a general framework in which to construct theories of specific interactions captures a central feature of modern theoretical physics. Newtonian theory, special relativity, and Einsteinian gravitation are *all* what might be called 'framework-theories.'[65] That is to say, they all provide frameworks of constraints in which physical quantities can be constructed and whose evolution can be determined. As Einstein has put it, they are based on 'general characteristics of natural processes, principles from which mathematically formulated criteria are developed, which the various processes or the theoretical representations of them have to satisfy' (Einstein, 1919 [2002], p. 213). These theories must be presupposed for further theorising; namely, for the construction of theories of special systems, the theory of a point-particle or that of a perfect fluid, for example. So, while the spatio-temporal frameworks given by Newtonian theory, special relativity, and Einsteinian gravitation are not a priori in any Kantian sense, they are prior in this special sense. Nonetheless, the presupposition of these spatio-temporal frameworks does not preclude the possibility that some new theory will motivate their replacement. Einstein appreciated this and noted their foundation in empirical generalisations.

[65]The notion of a framework-theory can be prised from Einstein's (1919 [2002]) notion of a principle-theory. For details, see Flores (1999) and DiSalle (2006; 2012) who have argued convincingly for this interpretation.

Regarding Newtonian theory, special relativity, and Einsteinian gravitation as framework-theories in this sense reinforces that, in spite of the significant differences in their accounts of space-time structure, they are theories of the same kind. From this point of view, the accounts I have considered in 'There is no conspiracy' and 'On identifying background-structure,' distinguish artificially between Newtonian theory and special relativity, on the one hand, and Einsteinian gravitation, on the other. I hope to have shown that, though Einstein's geometrisation of gravitation was a significant advance, the distinctions between conspiratorial theories and non-conspiratorial theories, background-dependent theories and background-independent theories, do not arise from inherent differences among the theories but from a meta-theoretical or metaphysical basis. The Newtonian and special-relativistic inertial frame concepts, and the geometric objects whose representation they require, are not asserted a priori but are tied to empirical criteria; namely, the laws of motion, and a criterion for identifying time of occurrence. In the same way, the inertial frame concept peculiar to Einsteinian gravitation is tied to an empirical criterion, in this case a criterion for identifying two previously distinct kinds of motions.

5.2. Future work: The proper methodological analysis of the equivalence principle

While these essays have examined the Newtonian and 1905 inertial frame concepts and the critical analysis that was motivated by the equivalence principle, there is still more to say about the methodological analysis of the principle. This represents a natural extension of this project.

The equivalence principle as a criterion of identity

Most work on the equivalence principle has focused on challenges that arise in the formulation of a statement of the principle and on the proper understanding of its scope of applicability (e.g., Pauli, 1921 [1958]; Anderson & Gautreau, 1969). Other work focuses on conceptual tangles that the principle supposedly raises (e.g., Eddington, 1923; Synge, 1960 [1971]; Ohanian, 1977; Norton, 1985). Still other work (Okon & Callendar, 2011) examines the equivalence principle with an eye to quantum theory.

This work is important. But it largely neglects the *methodological analysis* of the equivalence principle. A methodological analysis must consider two basic questions: What kind of principle is the equivalence principle? What is its role in the conceptual framework of gravitation theory? Part of the reason why a proper methodological analysis has not been pursued is because the equivalence principle has been characterised as a mere heuristic. We find such a view in Synge (1960 [1971], pp. ix-x), along with a dismissal: 'I have never been able to understand this principle .... The principle performed the essential office of midwife at the birth of general relativity .... I suggest that the midwife be now buried with appropriate honours.' In this way it has been relegated to part of the motivational basis for Einstein's construction of his gravitation theory. It is thought to have motivated Einstein's construction of his theory in much the same way Mach's principle did. But, while it did motivate the construction of a new gravitation theory, its role is not ambiguous in the way that Mach's principle was; its role was decisive.

The methodological function of the equivalence principle in the conceptual framework of physics appears in each essay. In 'Friedman's Thesis,' I addressed the principle directly. (In 'There is no conspiracy' and 'On identifying background-structure,' I was concerned with the equivalence principle insofar as it is part of the basis that motivates the application of a geometry of variable curvature to the world of experience.) Friedman's analysis consists of answers to the questions: What kind of principle is the equivalence principle? What is its role in the conceptual framework of physics? But, while I am sympathetic to Friedman's approach to the analysis of physical theories, my main task in 'Friedman's Thesis' is critical.

In future work, I would like to develop a positive account of the equivalence principle. The idea I want to develop further is that the equivalence principle is not about equivalence—in the sense of proportionality or behavioural indistinguishability—at all but about the recognition that two previously distinct concepts of motion are identical.

The idea to be developed further is that the equivalence principle is properly understood as a *criterion of identity*.

The notion of a criterion of identity has its origin in Frege (1884 [1980]), who made it the cornerstone of his theory of number. Frege sought to provide an analysis of the nature of arithmetic by showing that the theory of the natural numbers can be derived from a principle that has the same scope and generality as conceptual thought itself. This principle has the form: 'For any concepts *F* and *G*, the number of *F*s is the same as the number of *G*s if, and only if, there is a one-to-one correspondence between the *F*s and the *G*s.' (Demopoulos, 2013, p. 19) Frege introduced this principle as a criterion for assessing the conditions under which we should judge that the same number has been presented to us in two different ways as the number of two different concepts. In the context of second-order logic, Frege's criterion of identity implies the Peano-Dedekind axioms. But its role does not end there: The criterion also governs our judgements of equinumerosity in our *applications* of the theory of natural numbers; for example, in our everyday application of the theory of the natural numbers in counting.

Demopoulos (2013) has argued that the notion of a criterion of identity has a role to play in the methodological analysis of the exact sciences well beyond the one it plays in Frege's theory. Frege's notion of a criterion of identity can form the basis for a new account of the application of mathematical theories: Just as Frege showed that his criterion of identity for number governs our application of the theory of the natural numbers in counting, other criteria of identity govern the application of other mathematical theories; among them, Euclidean and Minkowskian geometry.

I want to show that the equivalence principle consists in the provision of a *criterion for identifying the motion of a classical inertial frame with that of locally freely falling one*. Furthermore, the provision of this criterion of identity is part of the basis that governs the application of the geometry of variable curvature in Einsteinian gravitation.

By developing this new account of the equivalence principle, I hope to deepen our understanding of the foundations of Einsteinian gravitation. But the account is also of more general importance to the history and philosophy of science. It offers a more nuanced and preferable alternative to two prominent accounts of scientific theory change, namely, Kuhn's (1962 [1970]) and the conventionalists' (e.g., Reichenbach, 1928 [1958]; Carnap, 1934 [1951]). These accounts have no better mechanisms to explain the transition from the 1905 inertial frame concept to the 1907 concept than the problematic notions of a Kuhnian paradigm shift and a change of conventions. On my account, this transition emerges as the result of an *analysis*, which avoids the pitfalls of Kuhn and conventionalism. Still another contribution of my account is that it substantiates the idea that criteria of identity have a wider role to play in methodological analysis.

The distinctive feature of Einstein's contribution

Having addressed the methodological status of the equivalence principle, I will isolate Einstein's distinctive contribution to our understanding of it. Isolating that contribution is important because the idea that inertial frames and freely falling frames are indistinguishable was in fact anticipated by Newton. Newton points out in Corollary VI to the laws of motion that if bodies moving with respect to one another are influenced by uniform accelerative forces along parallel lines they will move with respect to one another in the same way they would if they were not influenced by those forces. In other words, matter obeys the same laws in a freely falling frame that it would in an inertial frame.

The significance of Corollary VI is a subject of debate in the post-Einsteinian context. To begin with, it is controversial what the significance of Corollary VI is for the correct mathematical setting of Newtonian theory. Saunders (2013) claims that when Corollary VI is properly understood Newton's theory requires a mathematical setting different from those previously proposed. Saunders' claim is surprising; it goes against received wisdom about the correct mathematical setting of Newtonian theory. I want to evaluate Saunders' claim and compare his mathematical setting of the theory with a few of the usual candidates, including Galilean space-time, Maxwellian space-time, and geometrised Newtonian gravitation. Secondly, I want to argue that, though Corollary VI establishes that inertial motion and freely falling motion are indistinguishable, its function is not to identify them; that is, *Corollary VI is not a criterion of identity*. One could say that Newton failed to recognise freely falling frames and inertial frames as different presentations of the same concept. But, given his own theoretical framework, Newton was correct to think that freely falling frames and inertial frames are only equivalent to the degree that the freely falling system is very far from the source of the gravitational field, and so the lines of uniform acceleration are very nearly parallel; that is, to the degree that inhomogeneities in the gravitational fields external to the system in question are negligible. Newton demands only a 'near enough' isolated system for the sake of getting a 'near enough' measure of the forces at work within the system.

This work will contribute to the foundations of both Newton's and Einstein's gravitation theories. Identifying the correct mathematical setting is essential for correctly understanding Newtonian theory. By challenging the misconception that the main insight of the equivalence principle was already there in Corollary VI, this work will clarify what is distinctive about Einstein's contribution. Corollary VI and the 1907 equivalence principle are two interpretations, within different frameworks, of the same fundamental fact about the behaviour of bodies in gravitational fields.

The methodological status of background-independence

In 'On identifying background-structure,' I asked: What is the background-structure that Smolin would have us identify and remove? I took up the task of identifying three candidates for background-independence, and I argued that background-independence is something of a chimaera: No sooner do we seem to be getting a hold on the concept than it slips away again.

In future work, I want to examine the methodological status of Smolin's new heuristic principle: 'Seek to make progress by identifying the background-structure in our theories and removing it.' (Smolin, 2006, p. 204) Smolin's commitment to that principle is motivated by a number of metaphysical principles that he regards as genuine insights about nature, and that he takes to be satisfied by Einstein's gravitation theory. One example of such a principle is the relationist thesis of Leibniz (1715-16 [1970]) and Mach (1883 [1919]), according to which space and time are not real entities but an abstraction from the geometrical relations among bodies. Another is Leibniz's action-reaction principle. I think it is important to show that, whatever the heuristic value of such principles in theory construction, Einsteinian gravitation, at least, owes little or nothing to them. Take, for example, the relationist thesis, which Leibniz intended as a criticism of Newton's theory. The thesis fails doubly: It fails as a criticism of Newton's theory precisely because Leibniz did *not* postulate a kind of background-structure that Newton showed to be essential for the construction of an empirically adequate theory of motion; and it fails because the analogy commonly drawn between Leibniz's account of motion and Einsteinian gravitation is a bad one. In another line of argument, I want to show that there is something peculiar about Einstein's geometrisation of gravity. Although that geometrisation was an empirically well-motivated move in the construction of his gravitation theory—one closely knit to the equivalence principle—there is not an equally clear motivation to formulate the particular character of that theory as a much more general kind of methodological principle.

This continuation of my work on background-independence aims to further undermine the concept. But its contribution is not uniquely negative: It will illuminate what is distinctive about Einstein's geometrisation of gravitation. More generally, this work is of importance because it exposes an old and persistent pattern of reasoning about theories, one that gives undue importance to purely philosophical principles that are alleged to motivate genuine physical insights. In this respect, my work urges a stricter empiricism.

References

Anderson, J. L. (1967). *Principles of Relativity Physics*. New York: Academic Press.

Anderson, J. L. & R. Gautreau. (1969). Operational Formulation of the Principle of Equivalence. *Physical Review, 185*, 1656-1661.

Carnap, R. (1951). *The Logical Syntax of Language* (A. Smeaton, Trans.). London: Routledge & Kegan Paul. (Original work published in 1934)

Demopoulos, W. (2013). *Logicism and its Philosophical Legacy*. Cambridge: Cambridge University Press.

DiSalle, R. (2012). Analysis and Interpretation in the Philosophy of Modern Physics. In M. Frappier, D. Brown, & R. DiSalle (Eds.), *Analysis and Interpretation in the Exact Sciences: Essays in Honour of William Demopoulos* (pp. 1-18). Dordrecht: Springer Netherlands.

DiSalle, R. (2006). *Understanding Space-time*. Cambridge: Cambridge University Press.

Eddington, A. (1923). *The Mathematical Theory of Relativity*. Cambridge: Cambridge University Press.

Einstein, A. (2002). What is the theory of relativity? In M. Janssen & al. (Eds.), *The Collected Papers of Albert Einstein, vol. 7* (pp. 206-215). Princeton: Princeton University Press. (Original work published in 1919)

Flores, F. (1999). Einstein's theory of theories and types of theoretical explanation. *International Studies in the Philosophy of Science, 13*, 123-134.

Frege, G. (1980). *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number* (J. L. Austin, Trans.). Evanston: Northwestern University Press. (Original work published in 1884)

Gupta, A. (2006). *Empiricism and Experience*. Oxford: Oxford University Press.

Kuhn, T. (1970). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press. (Original work published in 1962)

Leibniz, G. (1970). *Philosophical Papers and Letters.* Ed. L. Loemker. Chicago: University of Chicago Press. (Original work published in 1715-16)

Mach, E. (1919). *The Science of Mechanics: A Critical and Historical Account of its Development* (T. McCormack, Trans.). Chicago: Open Court. (Original work published in 1883)

Norton, J. (1985). What was Einstein's Principle of Equivalence? *Studies in the History and Philosophy of Science, 16*, 5-47.

Ohanian, H. (1977). What is the Principle of Equivalence? *American Journal of Physics*, 45, 903-909.

Ohanian, H. (1976). *Gravitation and Spacetime*. New York: W. W. Norton & Company.

Okon, E. & C. Callender. (2011). Does quantum mechanics clash with the equivalence principle—and does it matter? *European Journal for Philosophy of Science, 1*, 133-145.

Pauli, W. (1958). *Theory of Relativity* (G. Field, Trans.). London: Pergamon. (Original work published in 1921)

Reichenbach, H. (1958). *The Philosophy of Space and Time* (M. Reichenbach & J. Freund, Trans.). New York: Dover. (Original work published in 1928)

Saunders, S. (2013). Rethinking Newton's *Principia. Philosophy of Science*, *80*, 22-48.

Smolin, L. (2006). The Case for Background Independence. In D. Rickles, S. French, & J. Saatsi (Eds.), *The Structural Foundations of Quantum Gravity*. Oxford: Oxford University Press.

Synge, J. (1971). *Relativity: The General Theory*. Amsterdam: North-Holland Publishing Company. (Original work published in 1960)

Appendix A

Formulating a principle of momentum conservation in Newtonian mechanics and special relativity

In this appendix, I review how to formulate a principle of conservation of momentum for a Newtonian or a special-relativistic system. This way of developing the conservation principle can be found in most elementary mechanics texts. I have followed Knight's *Physics* (San Francisco: Addison Wesley, 2004).

Consider a system of N particles. Identify each particle by a number k. Every particle interacts with every other particle via action-reaction pairs of forces. Furthermore, every particle may be subject to forces from outside the system.

Define the total momentum **P** of the system as the vector sum:

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + \cdots + \mathbf{p}_n = \sum_{k=1}^{N} \mathbf{p}_n.$$

The time derivative of **P** tells us how the total momentum of the system changes:

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{dt}} = \sum_{k} \frac{\mathrm{d}\mathbf{p}_k}{\mathrm{dt}} = \sum_{k} \mathbf{F}_k.$$

The net force acting on particle k can be divided into interaction forces within the system and external forces:

$$\mathbf{F}_k = \sum_{j \neq k} \mathbf{F}_{j \text{ on } k} + \mathbf{F}_{ext \text{ on } k}.$$

(Note that the restriction $j \neq k$ expresses the fact that particle k does not exert a force on itself.) Now we can formulate the rate of change of the total momentum **P** of the system:

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{dt}} = \sum_{k} \sum_{j \neq k} \mathbf{F}_{j \text{ on } k} + \sum_{k} \mathbf{F}_{ext \text{ on } k}.$$

The double sum adds the interaction forces within the system. But because of the third law $\mathbf{F}_{k \text{ on } j} = -\mathbf{F}_{j \text{ on } k}$ and $\mathbf{F}_{k \text{ on } j} + \mathbf{F}_{j \text{ on } k} = 0$. Thus, the sum of all the interaction forces is 0. That is to say, they form an equilibrating set. Therefore, the above equation becomes

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{dt}} = \sum_k \mathbf{F}_{ext \text{ on } k} = \mathbf{F}_{net}.$$

Here, $\mathbf{F}_{net}$ is the net force exerted on the system from without. *Note that this is just the second law written for the system as a whole*; that is, the rate of change of the total momentum of the whole is equal to the net force applied to the whole system.

For an isolated system, the above equation becomes

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{dt}} = 0.$$

In this way, we obtain a statement of the principle of conservation of momentum. *But note that this is just another way of expressing the principle of inertia.*

Appendix B

The notion of a constitutive principle in Helmholtz and Poincaré

While the laws of motion are the paradigm-example of a set of constitutive principles, examples can be found elsewhere in physics. Helmholtz's principles that lie at the basis of the geometries of constant curvature are also examples of constitutive principles. Helmholtz's principles originate in his analysis (1868 [1977a]; 1868 [1977b]; 1869 [1977c]) of Kant's theory of spatial intuition, according to which the Euclidean character of space is reflected in our ability to carry out the straightedge-and-compass constructions in *Elements*. Helmholtz recognised that the fundamental notions underlying these constructions are congruence and straightness, and he asked whether they have an empirical origin. The fruit of this analysis is his reply that our idea of space is neither innate nor a transcendental form of intuition but arises from our experiences of bodily motion and vision.

We learn through bodily motion that our bodies are not deformed as we move through space. That is, we infer it from the fact that there are no sensations specifically related to movement. From this experience, and from our experience with rigid bodies, we form a conception of the structure of space that is grounded in the study of displacements. The group of displacements that we discover in this way is summarised in the principle of free mobility: A rigid body may undergo arbitrary continuous motions without change of shape or dimension. The principle is constitutive of our concept of congruence, and so is one of the conceptual prerequisites for carrying out straightedge-and-compass constructions and, more generally, for making a measurement of length using a pair of dividers, a measuring rod or a chain.

We learn by reaching for objects in our visual field and by shielding our eyes from light that we can make certain judgements of direction and distance, and in this way we learn that the paths of light rays may be treated as straight lines. This is summarised in the principle of the straight-line propagation of light: The path of a light ray, inferred from our experience with lines of sight, may be treated as a straight line. The principle controls the application of the Euclidean concept of straightness; it expresses an empirical criterion by which straightness can be determined.[66] It is noteworthy that this does not imply that people without sight cannot have a concept of straightness; what perhaps they cannot have is the Kantian intuition of the construction of straight lines in visual imagination.

Helmholtz's analysis revealed that these principles control the application of Euclidean geometry. But his analysis revealed something further: The principle of free mobility is enough[67] to derive the Euclidean form of the metric.[68] The metric derivable from free mobility, however, admits three different isometry groups, that compatible with Euclidean geometry and also those compatible with the geometries of constant positive and negative curvature. The principle of free mobility restricts the mathematically possible spaces—all those compatible with a Riemannian manifold—to those of constant curvature. What is remarkable is that Helmholtz's analysis reveals that the same principle, free mobility, underlies both our intuitive conception of space and the mathematical conception of a class of metrical spaces.

With this important result, the question arises: Which of the three geometries of constant curvature is the correct geometry for the space of experience? Helmholtz replied that Euclidean geometry is singled out by our investigation of the behaviour of rigid bodies and our subsequent discovery that measurements are found to satisfy the axioms of Euclidean geometry.

The nature of Helmholtz's principles was further clarified by Poincaré (1902 [1952]). For both Poincaré and Helmholtz, the principle of free mobility and the principle of the straight-line propagation of light are constitutive of the geometries of constant

---

[66]Helmholtz argued that we can nonetheless imagine that space is non-Euclidean by picturing light rays that behave like the straight lines of a non-Euclidean space, that is, by imagining what visual sensations we would have. He even went so far as to suggest how we might create such sensations using lenses.
[67]It is unsurprising for this reason that the principle of free mobility figures much more prominently than the principle of straight-line propagation in Helmholtz's account.
[68]This result was only sketched by Helmholtz; it was proved by Lie in his theory of continuous groups. For this reason, the result is generally referred to as the Helmholtz-Lie theorem. See Stein (1977) and Torretti (1978) for details.

curvature. But, while Helmholtz held that Euclidean geometry is singled out by our investigation of the behaviour of rigid bodies and by our measurement of the angles of triangles, Poincaré disagreed. Poincaré held that, though experience furnishes us with the idea of a group of free motions, experience does not single out one of the three geometries of constant curvature. For him, there is no fact of the matter about which of the three geometries is the actual space of experience. If, however, we want to construct a dynamical theory, we must stipulate one of the geometries as a background; and, since the laws of such a theory will be simplest on a Euclidean background, Poincaré held that it would always be preferred.

The difference between Helmholtz's and Poincaré's views hinges on their analyses of the status of the principles; specifically, on the question whether they express facts or definitions. That there is a problem in holding the principle of free mobility to be an empirical fact was recognised by Helmholtz, who saw that free rigid motion and congruence are interdefinable: 'we have no criterion for the fixity of bodies and spatial structures other than that when applied to one another at any time, in any place and after any rotation, they always show the same congruences as before' (Helmholtz, 1868 [1977a], p. 24). With this recognition, one might say that for Helmholtz, then, the real empirical fact is the existence of a set of *relatively* rigid bodies, since, like Poincaré, he also appreciated that actual rigidity is an idealisation. But only in the work of Poincaré was the problem made explicit and resolved. While, for Helmholtz, the principle of free mobility and the principle of the straight-line propagation of light are the empirical 'facts in perception,' for Poincaré, they are 'definitions in disguise': The invariance of certain bodies under free motions is not a fact about those bodies but a definition of congruence; the straight-line propagation of light rays is not a fact about light but the definition of a straight line. Poincaré's philosophical contribution consists therefore in a correction; namely, in the recognition that Helmholtz had mistaken definitions for empirical hypotheses.

As in the case of Newton's laws, Helmholtz's principles are constitutive in the sense that they interpret the concepts of congruence, rigidity, and straightness by expressing criteria for their application. In this way, the principles control the application of the geometries of constant curvature.

Another example of a set of constitutive principles is found in Maxwell's equations. Maxwell's equations, like the laws of motion, appear to inform us about theoretical concepts that in fact they define and interpret for us. Much as in the Newtonian context, the basic materials of Maxwell's theory were already known pre-systematically from experience with permanent magnets and charged objects, from the discoveries of Oersted, Ampère, Faraday, and many others. But they remained only a piecemeal collection of facts and rules of thumb until Maxwell synthesised and reinterpreted them to construct a coherent system of scientific knowledge. Maxwell's equations constitute the concepts of electric field and magnetic field by expressing criteria for their application; they determine a framework of empirical investigation in which the relevant quantities can be detected and measured. Smith (2010) points out, for example, that the Ampère-Maxwell law is presupposed to a high degree of precision every time current is measured using a galvanometer of whatever sort.

References

Helmholtz, H. (1977a). On the Origin and Significance of the Axioms of Geometry. In R. Cohen & Y. Elkana (Eds.), *Hermann von Helmholtz: Epistemological Writings. The Paul Hertz/Moritz Schlick Centenary Edition of 1921* (pp. 1-38). Dordrecht: D Reidel Publishing Company. (Original work published in 1868)

Helmholtz, H. (1977b). On the Facts Underlying Geometry. In R. Cohen & Y. Elkana (Eds.), *Hermann von Helmholtz: Epistemological Writings. The Paul Hertz/Moritz Schlick Centenary Edition of 1921* (pp. 39-71). Dordrecht: D Reidel Publishing Company. (Original work published in 1868)

Helmholtz, H. (1977c). The Facts in Perception. In R. Cohen & Y. Elkana (Eds.), *Hermann von Helmholtz: Epistemological Writings. The Paul Hertz/Moritz Schlick Centenary Edition of 1921* (pp. 115-185). Dordrecht: D Reidel Publishing Company. (Original work published in 1869)

Poincaré, H. (1902). *Science and Hypothesis* (W. Greenstreet, Trans.). New York: Dover. (Original work published in 1952)

Smith, G. (2010). Revisiting accepted science: The indispensability of the history of science. *The Monist, 93*, 545-579.

Stein, H. (1977). Some Philosophical Prehistory of General Relativity. In J. Earman, C. Glymour, & J. Stachel (Eds.), *Foundations of Space-Time Theories. Minnesota Studies in the Philosophy of Science, vol. VIII* (pp. 3-49). Minnesota: University of Minnesota Press.

Torretti, R. (1978). *Philosophy of Geometry from Riemann to Poincaré*. Dordrecht-Holland: D. Reidel Pub. Co.

VITA

Ryan Samaroo

Biographical Data

1. Personal
Born 1981, Toronto

2. Degrees
BA, Université Laval, 2005
PhD, University of Western Ontario, 2013

Research

3. Research and teaching interests
a. Areas of specialisation
History and philosophy of science, especially the history and philosophy of physics from Newton to the present; connections between philosophy of science and analytic philosophy

b. Areas of competence
History and philosophy of mathematics, especially the foundations of geometry and arithmetic; history of modern philosophy, from Descartes through logical empiricism

4. Publications
'On identifying background-structure in classical field theories,' *Philosophy of Science* 78(2011): 1070-1081.

5. Significant graduate-level awards
Social Sciences and Humanities Research Council of Canada
Doctoral Scholarship, 2008-2010

Canadian Society for the History and Philosophy of Science
Richard Hadden Prize, 2008