PHILOSOPHY and
FOUNDATIONS
of PHYSICS
Vol. 4

Series Editors: D. Dieks and M. Redei

The Ontology of Spacetime II

Dennis Dieks
(Editor)

Philosophy and Foundations of Physics

Series Editors: Dennis Dieks and Miklos Redei

In this series:

Vol. 1: The Ontology of Spacetime
Edited by Dennis Dieks

Vol. 2: The Structure and Interpretation of the Standard Model
By Gordon McCabe

Vol. 3: Symmetry, Structure, and Spacetime
By Dean Rickles

Vol. 4: The Ontology of Spacetime II
Edited by Dennis Dieks

The Ontology of Spacetime II

Edited by

Dennis Dieks

Institute for History and Foundations of Science
Utrecht University
Utrecht, The Netherlands

ELSEVIER

Amsterdam – Boston – Heidelberg – London – New York – Oxford – Paris
San Diego – San Francisco – Singapore – Sydney – Tokyo

Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

First edition 2008

Copyright © 2008 Elsevier B.V. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system
or transmitted in any form or by any means electronic, mechanical, photocopying,
recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights
Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333;
email: permissions@elsevier.com. Alternatively you can submit your request online by
visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting
Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons
or property as a matter of products liability, negligence or otherwise, or from any use
or operation of any methods, products, instructions or ideas contained in the material
herein. Because of rapid advances in the medical sciences, in particular, independent
verification of diagnoses and drug dosages should be made

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN-13: 978-0-444-53275-6
Series ISSN: 1871-1774

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in Hungary

08 09 10 11 12 10 9 8 7 6 5 4 3 2 1

CONTENTS

List of Contributors vii

Preface ix

1. A Trope-Bundle Ontology for Field Theory **1**

2. Is Structural Spacetime Realism Relationism in Disguise? The Supererogatory Nature of the Substantivalism/Relationism Debate **17**

3. Identity, Spacetime, and Cosmology **39**

4. Persistence and Multilocation in Spacetime **59**

5. Is Spacetime a Gravitational Field? **83**

6. Structural Aspects of Space-Time Singularities **111**

7. Who's Afraid of Background Independence? **133**

8. Understanding Indeterminism **153**

9. Conventionality of Simultaneity and Reality **175**

10. Pruning Some Branches from "Branching Spacetimes" **187**

11. Time Lapse and the Degeneracy of Time: Gödel, Proper Time and Becoming in Relativity Theory **207**

12. On Temporal Becoming, Relativity, and Quantum Mechanics **229**

13. Relativity, the Passage of Time and the Cosmic Clock **245**

14. Time and Relation in Relativity and Quantum Gravity: From Time to Processes **255**

15. Mechanisms of Unification in Kaluza–Klein Theory **275**

16. Condensed Matter Physics and the Nature of Spacetime **301**

Subject Index 331
Author Index 337

LIST OF CONTRIBUTORS

Richard T.W. Arthur, Department of Philosophy, McMaster University, Hamilton, Canada

Jonathan Bain, Humanities and Social Sciences, Polytechnic University, Brooklyn, USA

Yuri Balashov, Department of Philosophy, University of Georgia, Athens, USA

Tomasz Bigaj, Institute of Philosophy, Warsaw University, Warsaw, Poland

Carolyn Brighouse, Department of Philosophy, Occidental College, Los Angeles, USA

Mauro Dorato, Department of Philosophy, University of Rome 3, Rome, Italy

John Earman, Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, USA

Jan Faye, Department of Media, Cognition, and Communication, University of Copenhagen, Copenhagen, Denmark

Peter Forrest, School of Social Science, University of New England, Armidale, Australia

Vincent Lam, Department of Philosophy and Centre Romand for Logic, History and Philosophy of Science, University of Lausanne, Lausanne, Switzerland

Dennis Lehmkuhl, Oriel College, Oxford University, Oxford, UK

Ioan Muntean, Department of Philosophy, University of California, San Diego, USA

Vesselin Petkov, Department of Philosophy, Concordia University, Montreal, Canada

Dean Rickles, History and Philosophy of Science, The University of Sydney, Sydney, Australia

Alexis de Saint-Ours, University of Paris VIII and Laboratory "Pensée des Sciences", École Normale Supérieure, Paris, France

Andrew Wayne, Department of Philosophy, University of Guelph, Canada

PREFACE

The sixteen papers collected in this volume are expanded and revised versions of talks delivered at the Second International Conference on the Ontology of Spacetime, organized by the International Society for the Advanced Study of Spacetime (John Earman, President) at Concordia University (Montreal) from 9 to 11 June 2006.

In the First Conference, held in 2004, the majority of the papers were devoted to topics relating to *Becoming* and the *Flow of Time*.[1] Although this subject is still well represented in the present volume, it has become less dominant. Most papers are now devoted to subjects directly relating to the title of the conference: the *ontology of spacetime*.

The book starts with four papers that discuss the ontological status of spacetime and the processes occurring in it from a point of view that is first of all conceptual and philosophical. The focus then slightly shifts in the five papers that follow, to considerations more directly involving technical considerations from relativity theory. After this, *Time, Becoming* and *Change* take centre stage in the next five papers. The book ends with two excursions into relatively uncharted territory: a consideration of the status of Kaluza–Klein theory, and an investigation of possible relations between the nature of spacetime and condensed matter physics, respectively.

The marked differences between the programs of the First and the Second Conference, respectively, and the large audiences assembled on both occasions, bear witness to the vitality of the field of Philosophy and Foundations of Spacetime. Preparations for the Third Conference, in 2008, are already well on their way!

Dennis Dieks
History and Foundations of Science,
Utrecht University, Utrecht,
The Netherlands

[1] See: Dieks, D. (Ed.), 2006. The Ontology of Spacetime. Elsevier, Amsterdam.

CHAPTER 1

A Trope-Bundle Ontology for Field Theory

Andrew Wayne*

Field theories have been central to physics over the last 150 years, and there are several theories in contemporary physics in which physical fields play key causal and explanatory roles. This chapter proposes a novel field trope-bundle (FTB) ontology on which fields are composed of bundles of particularized property instances, called tropes (Section 2) and goes on to describe some virtues of this ontology (Section 3). It begins with a critical examination of the dominant view about the ontology of fields, that fields are properties of a substantial substratum (Section 1).

1. FIELDS AS PROPERTIES OF A SUBSTANTIAL SUBSTRATUM

The dominant view about the ontology of field theory over the last two centuries has been that fields are properties of a substantial substratum. In the 19th century this substance was taken to be a material ether. In the 20th century, the immaterial spacetime manifold took on the role of substantial substratum.

For most of the 19th century, the causal and explanatory functions of field theories were assumed by a material, mechanical ether. Field theories of optics, electricity, magnetism and later electromagnetism were developed in which the field corresponded to a collection of properties of a material ether. Scientists articulated the hope that a unified theory could be extended to gravitational and other phenomena, where a single material ether would be the seat of all physical action. George Green and Lord Kelvin, for example, developed optical theories in which light was the vibration of a mechanical, elastic, solid ether (Green, 1838; Kelvin, 1904). This ether was made up of tiny ether particles. Lagrangian mechanics, augmented with a few auxiliary hypotheses, were used to obtain many

* Department of Philosophy, University of Guelph, Guelph, ON, N1G 2W1, Canada
E-mail: awayne@uoguelph.ca

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00001-6
© Elsevier BV
All rights reserved

sophisticated optical results: derivation of Fresnel's laws of reflection and refraction of light, phase shifts on reflection and elliptical polarization. From the start, however, these theories were extremely complex and ultimately only able to account for a narrow range of optical phenomena. As they were extended to new domains, ad hoc hypotheses were needed to make them work. For example, the value of the ether's resistance to distortion (shearing) needed to be set at one value to account for double refraction and another to account for Fresnel's laws.

Yet none of these difficulties was seen to impugn the mechanical ether hypothesis itself. The approach was extended to Maxwell's unified dynamical theory of light, electric and magnetic phenomena. Thus in the 1890s Joseph Larmor developed a sophisticated theory in which the ether is a kind of primitive continuous matter or proto-matter to which Maxwell's equations apply (Larmor, 1900). The electromagnetic field consists of undulations of this ether and electrons are singularities in the ether. The dynamics of ordinary matter are caused by the proto-material ether. Larmor and others around the turn of the century understood the materiality of the ether to amount to the fact that it has mechanical properties and can engage in mechanical interactions. Larmor's account ran into difficulties, and some of these difficulties were taken to be endemic to any material ether theory. No one was able to develop an empirically adequate theory of electrodynamic phenomena based on the principle of least action and the interaction between matter and a proto-material ether.

The most important response to this problem was H.A. Lorentz's theory in which the electromagnetic field consists of a collection of properties of an immaterial ether. Lorentz's ether functioned as a unique, immutable reference frame for electrodynamics. Lorentz explicitly rejected mechanical ether theories and adopted as his fundamental assumption "that ponderable matter is *absolutely* permeable [to the ether], i.e., that the atom and the ether exist in the same place" (Lorentz, 1895, Section 1). Matter has no effect on the ether, but the ether can causally affect matter, and the ether remains the seat of the electromagnetic field. In addition, the null result of the Michelson–Morley experiment was accounted for by the Lorentz–Fitzgerald contraction, itself taken to be directly caused by motion of matter with respect to the ether. Of course, no experiment was able to distinguish the rest frame of the ether. Worse, Einstein's highly successful 1905 special theory of relativity was taken to be inconsistent with the postulation of any privileged frame of reference. Fully aware of this, Lorentz still could not give up the ether. In his 1909 book *The Theory of Electrons* Lorentz offers a detailed account of the virtues of Einstein's approach, in the middle of which he remarks:

> Yet, I think, something may also be claimed in favour of the form in which I have presented the theory. I cannot but regard the ether, which can be the seat of the electromagnetic field with its energy and its vibrations, as endowed with a certain degree of substantiality, however different it may be from ordinary matter (Lorentz, 1909; quoted in Schaffner, 1972, p. 115).

Lorentz's intuition here seems to be that the only way the electromagnetic field can play the causal and explanatory roles it does is if the field is a substantial entity. This substantiality appeared to Lorentz to be secured by an immaterial ether.

19th-century field theories were formulated within the context of 19th-century metaphysics, and of course the dominant metaphysical posits of that century were the connected notions of substance and attribute. The notion of substance traditionally involves three elements. First and most intuitive is the idea that a substance is something that can have independent existence, whereas an attribute cannot but is rather a dependent entity. Second, substance plays the role of bearer of attributes: a substance has attributes inhering in it but need not itself inhere in anything. Third, substance functions to individuate one property from other, possibly exactly alike properties.

Field theories with a material ether ontology are the quintessential scientific articulation of a substance-attribute metaphysics. Here, a material ether is a substance and classical fields consist of properties (attributes) inhering in that substance; the ether is a sort of peg upon which field properties are hung. The notion of a material substratum is relatively straightforward, and within particular ether theories this substance is posited to have intrinsic properties of compressibility, resistance to shearing, and so on, independent of any additional, contingent attributes (such as field properties) it may have. The three traditional elements of the substance concept are well exemplified here. Clearly, a material ether can exist without any field, but the field cannot exist without the ether, giving the ether independent physical existence. As well, the ether bears properties, specifically the field properties. Finally, the ether functions to individuate field properties. Two exactly-alike field values are individuated and indexed by the ether, the substantial substratum in which they inhere. If there ever were a case for a traditional substance-attribute metaphysics, classical field theory would seem to be it.

It is more difficult to see how an immaterial ether, such as Lorentz's, can play the role of substantial substratum. For one thing, it is something of a mystery how an immaterial ether, absolutely permeable to material objects, can function as the bearer of a field, such as the electromagnetic field, that has a certain degree of materiality (it has energy and it causally interacts with ordinary matter). For another, the independent existence of the ether is mysterious, since it is simply posited to play the role of supporting the field, a seemingly ad hoc postulation. Third, there is the vexed question of whether the immaterial ether has essential properties in addition to the field or other accidental attributes it may bear. Lorentz's proposal seems to be that the immaterial ether has no essential properties, but rather is simply the "seat" of the field. An ether denuded of properties shares all the metaphysical troubles that face any bare particular. For example, points in the ether have an individuality, a haecceity, which enables them to be indexed and makes it possible for there to be more than one of them. But among the properties that bare particulars lack are any that would allow one to be distinguished from another. The implication is that there can at most be one point in the ether, or if there is more than one they can't be indexed. It appears that such an ether could play no useful role in the ontology of physical field theory. For these reasons we may be inclined to augment Lorentz's immaterial ether with certain geometrical properties, such as topological, differential and metrical properties, so that it can fulfill the role of indexer and individuator of field properties. This appears to be a promising

strategy and it is, moreover, precisely the direction taken by the ontology of 20th century field theories.

That fields are properties of a substantial substratum remains the received view to the present day. Now it is no longer an ether that is providing the substance, but rather the spacetime manifold. In contemporary physics, the spacetime manifold has replaced the ether as the substratum in which field properties inhere.

The ontology can be stated quite briefly: a field is an assignment of a collection of properties or field values (described by numbers, vectors or tensors) to points in spacetime. Field properties are causal properties and spacetime points function as independent causal agents in field theories, on a par with the causal agency of other physical objects. Spacetime points are necessary for field theory, since without them there is nothing to which field properties can be assigned and hence there can be no physical fields. As well, spacetime points are sufficient since no additional substance, matter or mechanism is needed. As Hartry Field puts it, "acceptance of a field theory is not acceptance of any extra ontology beyond spacetime and ordinary matter" (Field, 1989, 183; cf. Field, 1980, 35). John Earman describes the role of spacetime substance similarly:

> When relativity theory banished the ether, the spacetime manifold M began to function as a kind of dematerialized ether needed to support the fields.... [I]n postrelativity theory it seems that the electromagnetic field, and indeed all physical fields, must be construed as states of M. In a modern, pure field-theoretic physics, M functions as the basic substance, that is, the basic object of predication (Earman, 1989, 155).

On this approach, examples of classical fields that are properties of the spacetime substance include the metric field and stress-energy field of the general theory of relativity, and the electromagnetic field. Earman distinguishes first-order and second-order properties of spacetime points. First-order properties are the points' topological and differential properties, and field values constitute second-order properties.

We ought to question, however, whether the spacetime manifold, an immaterial ether with geometrical properties, can fulfil its role as the substantial substratum for classical field theories. For one thing, a worry raised earlier about the Lorentzian ether remains unresolved. Fields in contemporary physics are material objects; they contain mass-energy and interact causally with other material objects. At the very least, more needs to be said about how an immaterial spacetime, absolutely permeable to material objects, can function as the bearer of a material field. For another, the assumption that the spacetime manifold is a substance is controversial and faces a significant challenge from the hole argument of Earman and John Norton (Earman and Norton, 1987). *Cateris paribus*, it would seem preferable that a field ontology not be committed to spacetime substantivalism.

Perhaps an adequate ontology for classical fields could do without spacetime substance. David Malament has pointed out that the above characterization of fields as assignments of properties to points of spacetime can equally well (that is, poorly) be used to describe middle-sized material objects, such as his sofa.

> The important thing is that electromagnetic fields are "physical objects" in the straightforward sense that they are repositories of mass-energy. Instead of saying that spacetime points enter into causal interactions and explaining this in terms of the "electromagnetic properties" of those points, I would simply say that it is the electromagnetic field itself that enters into causal interactions (Malament, 1982, 532).

On this approach, field theories introduce a new kind of entity, fields, into our ontology. Fields have mass-energy, just like the kinds of physical entities with which we are more familiar, and they have additional properties unique to each field. Along similar lines, Paul Teller has proposed an inversion of the role of substance and attribute, this time in the context of a thorough-going relationalism about spacetime. Rather than attributing a field property to a spacetime point, he suggests attributing a relative spatio-temporal location to a bit of the substance making up the field (Teller, 1991, 382). Spatio-temporal relations are then carried by the field stuff directly.

It has been argued that, as it stands, the Teller position falls short of characterizing a genuine alternative. Hartry Field claims that if fields have all the geometric structure and causal powers that he attributes to spacetime, then there is no point in positing a separate, causally inert spacetime. Further, if we dispense with spacetime, as Teller does explicitly, the above response is trivialized: what Field calls "spacetime" Teller is simply calling "field," and the two approaches are equivalent (Field, 1989, 183). The same point has been made by Robert Rynasiewicz. Fields can be seen as properties of spacetime points, where the latter are construed as independently-existing individuals with specific additional (geometric) properties. Or fields can be viewed as collections of independently-existing individuals that have both causal (field) properties and the same geometric properties as did the spacetime points. These two pictures are ontologically equivalent; the difference between them is purely terminological, amounting to a disagreement over what should be called what (Rynasiewicz, 1996, 302–3; according to Rynasiewicz, Malament has acknowledged that his comments are intended to be read in this way). If this line of reasoning is correct, a consensus about the ontology of field theories in physics emerges, roughly that fields are properties of some substantial substratum, variously called the spacetime ether, spacetime manifold, or field stuff.

I suggest that this line of reasoning is not correct, and that Teller has articulated a genuine alternative—or at least that his approach is compatible with a very different ontological picture. The idea that the two approaches are equivalent may be plausible only if both approaches are formulated within the context of a substance-attribute ontology (although perhaps not even then: Belot (2000, 584) argues that this equivalence is implausible under certain relationalist assumptions). Taking a closer look at the roles that the substantial substratum plays in these approaches reveals an important difference between them. A substance-attribute ontology is indeed natural for the former view, on which fields are properties of a dematerialized ether. Here, the spacetime ether is clearly a substance, functioning to individuate and index the field attributes. This role for space is a traditional one. For instance, given two objects that are exactly alike, one knows they are two

objects, and not one, because they are located at different points in space. I suggest that on the latter view, where fields are independently-existing entities, the most natural ontology is one in which fields are composed of bundles of properties and relations. The "substance" making up the fields is nothing other than properties and relations. The Teller proposal is best understood within a pure property-bundle ontology, and it provides a genuine alternative to an ontology based on the spacetime manifold playing the role of immaterial substratum.

However, an ontology in which objects are composed of bundles of properties faces at least one well-known difficulty. This difficulty stems from the fact that properties are universals, where exactly alike properties of multiple objects are actually multiple instantiations of a single universal. The whiteness of this piece of paper and the exactly alike whiteness of that pen are strictly identical, since both objects instantiate the same universal, namely that shade of whiteness. So a property-bundle theory is committed to the necessary truth of the principle of the identity of indiscernibles. If an object is nothing more than a bundle of universals, then it is logically impossible for there to be two bundles with exactly the same properties. Two bundles composed of the same (universal) properties have all the same components, hence they would simply be the same bundle. However, it seems a contingent truth, if it is true at all, that distinct particulars must differ in their properties or relations (so the principle of the identity of indiscernibles is, if true, only contingently true). These difficulties are particularly acute in the case of field theory, where numerically distinct yet exactly alike point field values seem entirely plausible, as Earman has emphasized (1989, 197; cf. Parsons and McGivern, 2001).

2. FIELDS AS TROPE BUNDLES

Tropes are property instances, and they can be used to construct an ontology that is both nominalist, thus dispensing with universals, and bundle-theoretic, thus dispensing with the substantial substratum. It would seem that tropes are promising building-blocks for an ontology for field theories that can underwrite their causal and explanatory roles in contemporary physical theory. The remainder of this chapter attempts to make good on this promise.

Recall that exactly alike properties of multiple objects are multiple instantiations of a single universal. By contrast, exactly alike tropes of multiple objects are independent particulars. On this approach, the whiteness of this piece of paper and the exactly alike whiteness of that pen are numerically distinct tropes. An ontology in which objects are composed of bundles of tropes is not committed to the identity of indiscernibles. There is no special difficulty with having two bundles exactly alike, since each bundle contains its own particular tropes.

Trope-bundle ontologies face other worries, however. One challenge concerns the nature of the bundling relation that ties a collection of tropes together into an object. Tropes typically occur in *compresent* collections or bundles; for example, a patch of green paint can be analyzed as a collection of compresent tropes that

include, *inter alia*, a green trope, a *being at 18° C* trope and a place trope (this relation is called "concurrence" by Chris Daly (1994) and D.C. Williams (1997)). A key problem for trope ontologies is to give an account of this compresence relation. On the one hand, the compresence relation itself may be external to, and not founded upon, members of the collection of tropes that form its relata. Call this external relation $\text{compresence}_{\text{EX}}$. In this case the compresence relation does not supervene on the tropes it relates; it is an additional relational trope binding the two or more tropes in the bundle. On the other hand, compresence may be an internal relation, a consequence of the tropes themselves and not anything ontologically extra beyond the tropes in the bundle. Tropes bound by $\text{compresence}_{\text{IN}}$ are necessarily and essentially bundled. $\text{Compresence}_{\text{IN}}$ means that the independent particular—the bundle of tropes—cannot exist without each and every trope that composes it. A successful trope-bundle ontology needs to include a satisfactory account of compresence relations between bundled tropes.

A second worry about trope ontologies concerns how bundles of tropes, which are nothing more than instances of properties or relations, can play the substantial roles of bearing attributes and having independent existence. We have seen that the notion of substance involves three related ideas. One is the idea that a substance is something that can have independent existence, whereas an attribute cannot but is rather a dependent entity. Second, substance functions to individuate one property from other, possibly exactly alike properties (recall that the property-bundle approach ran into difficulty here). Third, substance is the bearer of attributes. A substance has attributes inhering in it but need not itself inhere in anything. Clearly, a successful trope-bundle ontology needs to show how these substantial roles are fulfilled by trope bundles.

It will be instructive to look at one well-known attempt to develop a trope ontology for field theory, that of Keith Campbell (1990). Campbell posits an ontology based exclusively on classical fields and spacetime. On his approach, a field, such as the electromagnetic field, pervades all spacetime. He is motivated to resist unfounded $\text{compresence}_{\text{EX}}$ relations because of what he sees as their derivative ontic status: "some tropes, the monadic ones, can stand on their own as Humean independent subsistents, while others, the polyadic [relational] ones are in an unavoidably dependent position" (1990, 99). The state of a field in four-dimensional spacetime is represented in the ontology by a *single* trope, and the field has no real, detachable parts. If more than one field exists, each one consists of a single trope. In the same way, all of spacetime itself corresponds to a single infinite, partless, edgeless trope (1990, 145–151). In this way, Campbell attempts to finesse the bundling problem by eschewing $\text{compresence}_{\text{EX}}$ relations entirely. Fields are essentially infinitely extended entities, he asserts, and if a field exists then it must *necessarily* be compresent with spacetime. Thus the compresence relation, in this case, is $\text{compresence}_{\text{IN}}$: it supervenes on, and is nothing ontologically over and above, the field trope itself (1990, 132–3).

As an account of the ontology of classical field theory, Campbell's proposal is unsatisfactory in several ways (cf. Moreland, 1997; Molnar and Mumford, 2003). A useful rule of thumb in analytic ontology is to avoid making substantive assumptions about how the world must be wherever these can be avoided. As

Campbell puts it, an adequate ontology "should leave open, as far as possible,… plainly *a posteriori* issues" (1990, 159). Yet Campbell's proposal is based on a number of very large such assumptions, some of which are not consistent with classical field theory. For example, it assumes that if a field exists it is necessarily coextensive with all spacetime; such an assumption is not consistent with classical field theory, as the latter is usually taken to allow for the physical possibility of null field values and so regions of spacetime in which the field is not present. Moreover, it seems to get the modalities wrong. On Campbell's approach, each and every occurrence of a trope in a bundle becomes a matter of necessity. However, we usually conceive of the world in terms of varying degrees of necessity. We want to distinguish, for instance, between those compresences of tropes that are necessary and those that are contingent.

In addition, Campbell's proposal is poorly motivated. The second-class ontic status Campbell imputes to relational (dyadic and polyadic) tropes comes from the fact they need to be borne by at least two other tropes, while "[m]onadic tropes require no bearer" (1990, 99). But most monadic tropes *do* require a bearer, or at least are dependent on one or more other tropes for their existence. A particular quality of greenness, a specific instance of *being at 18° C*, and so on, all require a complex of other tropes to sustain them and are thus equally in an "unavoidably dependent position." Even a classical field trope, as Campbell conceives it, depends upon a spacetime trope for its existence. There may be some lone tropes that can exist independently of the compresence of any other trope (Campbell's spacetime trope, for instance), but these are the exception rather than the rule. That dyadic and polyadic tropes require other tropes for existence does not distinguish them ontologically from monadic tropes, and is certainly no motivation for attempting to eliminate them from the ontology.

Campbell's proposal seems barely distinguishable from the very substance-attribute approach that trope theory is trying to do without. The field trope depends for its existence on a spacetime "peg," while the spacetime trope does not depend for its existence on any other trope. The spacetime trope performs the trick of augmenting the dependent particular (the field) in such a way that the pair becomes an independent particular. In short, the spacetime trope functions as a substantial substratum and, apart from its thoroughgoing eschewal of universals, Campbell's proposal amounts to a variant of a substance-attribute field ontology. We can do better.

The best place to begin an ontological assay of classical fields is with a characterization of what a field is. As it is usually described, a field consists of values of physical quantities associated with spacetime locations or spatiotemporal relations. We shall have more to say about what constitutes the "value of a physical quantity" below; for the moment think of intuitive values such as 0.3 Gauss of magnetic field strength. This central element of the field concept is, as a rule, given the following ontological gloss: in a field, values of a physical quantity inhere in and are properties of the ether or spacetime manifold. A field (a set of field values inhering in spacetime points) is thus a complex dependent particular that relies on a manifold or ether for its existence. We have explored this ontology and the

challenges it faces (Section 1). We shall now pursue an alternative ontology, the field trope-bundle (FTB).

The general structure of the field trope-bundle ontology is based on Peter Simons' "nuclear theory" (1994, 567–9). This ontology is characterized by kernels of $\text{compresent}_{\text{IN}}$ tropes that are themselves related by $\text{compresent}_{\text{EX}}$ relations. Simons is concerned exclusively with a trope-bundle ontology for particles and everyday objects. Our present task is to extend his approach to the case of physical fields.

The first step in the field trope-bundle construction identifies a kernel or core of tropes which must all be compresent. This kernel is necessary for a field to be a complex independent particular. The kernel at each point consist of three kinds of tropes, one or more G tropes, one or more F tropes, and an *x* trope. A G trope is a particular topological or metrical property instance of the spacetime at a point. F tropes are particular instances of field values (such as 0.3 Gauss magnetic field strength). Each G and F trope carries with it its own particularity, since *being a particular* is a basic fact about every trope. But particularity alone is not enough for G and F tropes to have independent existence. To see why, note that while particular entities can be aggregated, it is no part of the concept of particularity that particular entities must have numerical identity (i.e. can be indexed or labelled). Quantum-mechanical particles, for example, provide an example of particular entities that can be aggregated but do not have numerical identity (Redhead and Teller, 1991; Teller, 1995). The ontology of field theory, by contrast, requires that field values have a stronger individuality, one which supports indexing.

The complex of G and F tropes requires the *x* trope to index it. The *x* trope can be understood as a particular "way" that an G-F trope complex can be, namely one with a particular indexed identity. The *x* trope is thus not, by itself, substantial or substance-like and cannot exist without something else, the G-F trope complex, for it to be a way of. The collection of various *x* tropes has merely set-theoretic structure (ordinality and membership) and is not to be associated with a spacetime manifold. In Minkowski spacetime, for example, the G tropes are all exactly alike, while the *x* tropes each differ in their numerical identity. The kernel just described, consisting of a G-F trope complex and an *x* trope bound together, are the building blocks of fields.

We have been speaking of a G-F-*x* trope kernel at a point, but such talk may be inaccurate. F and G tropes may be best understood as irreducibly relational. Electromagnetic field values, for instance, can be understood as constituted by their counterfactual relations to other field values specified in the electromagnetic field equations, a hidden relationality. Geometrical property instances may also be understood as relational. This is accommodated naturally within the trope-bundle approach by accounting for relational property instances in terms of polyadic tropes that are compresent with more than one point field value (itself consisting of an *x* trope and any monadic tropes). Here field regions, rather than the point field values, are the basic independently-existing kernels.

The bundling relation *within* the kernel is $\text{compresence}_{\text{IN}}$: the compresence of the G-F-*x* tropes within a bundle supervenes on the tropes themselves. This is a consequence of the fact that within a kernel all tropes are necessarily compresent.

We also need to account for relations between, and compresence of, independent fields. Distinct fields consist of independently-existing field kernels. Two independent field kernels may be compresent$_{EX}$, where this sort of compresence is an external relation constituted by one or more relational tropes, called E tropes. If a field consists of more than one kernel, or if there is more than one independent field, then field kernels require some E tropes to be bundled with them, although which E trope or tropes is a contingent matter. In this way, E tropes are more loosely bound to the field kernels than are the tropes within the kernels themselves. Field kernels do not require specific E tropes in order to exist, and it is possible that the same (independently-existing) field kernel be part of different compresence$_{EX}$ relations, that is, be bound to different E tropes.

One virtue of the FTB ontology is that it responds, at least in part, to the two main challenges facing trope-bundle ontologies in general: the role of substance and the nature of the compresence relation. On the FTB proposal, each field point or, in the case of relational tropes, field region can be an independent particular. It is not that the field inheres in a substantial substratum, but rather that each field kernel *is* substantial. Recall that the notion of substance involves three related ideas. One is the idea that a substance is something that can have independent existence. This is true for field kernels as we have defined them (an alternative will be presented shortly). Another role of substance is that it functions to individuate one property from other, possibly exactly alike properties. Trope bundles in the FTB ontology fulfill that role, because tropes, as particulars, are automatically individuals, and the ontology contains a specific mechanism for rendering bundles numerically distinct. A third role for substance is as the bearer of attributes, and field kernels in the FTB construction play this role as well. Field kernels may function as the bearer of attributes by means of a compresence$_{EX}$ relation, where these attributes are also tropes. However, these particular attributes are not essential for the existence of the field kernel. This is in keeping with the ontological asymmetry between substance and attribute, where the substance exists independently but the attribute depends on the substance. It should be noted, however, that there remain significant difficulties in elucidating an external compresence relation for tropes (Simons, 1994; Daly, 1994).

A second virtue is that the FTB ontology is flexible and can accommodate the diversity of field theories in contemporary physics. This point is worth emphasizing, especially in light of the suggestion below that a trope-bundle approach might prove useful for analyzing ontological aspects of quantum field theory. One way a trope-bundle approach is flexible is with respect to what are the particular tropes that count as field values. So far we have referred to definite-valued field values, that is, field values that are determinate quantities of a physical variable (such as 0.3 Gauss magnetic field strength). It is worth noting that dispositions and propensities are equally tropes (Molnar and Mumford, 2003) and equally good candidates for field values. Another way a trope-bundle approach is flexible is with respect to the size of the field kernel. We began with the assumption that the field values plus geometrical property instances at a point constituted a kernel, and we then expanded the kernel to finite regions in order to include compresent$_{IN}$ relational

property instances as well. It may be that $\text{compresence}_{\text{IN}}$ is not limited to any finite region of the field, in which case the field as a whole is made up of a single kernel.

Another flexibility in the approach worth noting concerns whether the field kernel has independent existence. A kernel is a core of tropes which must all be compresent, and as we have seen, a field kernel can be an independent particular. This seems natural in the case of classical field theory. A system can contain an electromagnetic field and nothing else, for instance, so it is clear that the field can exist independently of anything else; kernels in the electromagnetic field are independently-existing entities. However, nothing in the FTB ontology requires that this be so. It may be the case that a field kernel cannot exist independently of a periphery of other tropes to which it is bound by $\text{compresence}_{\text{EX}}$ relations. This dependence between the kernel and the periphery would be token-type, so that tropes within the kernel depend on there being some trope $\text{compresent}_{\text{EX}}$ in the periphery of a certain type. In the same way, middle-sized physical objects must be compresent with some temperature trope, although they do not generally depend on any one specific temperature trope for their existence. The FTB ontology for quantum field theory sketched below provides an example of field kernels that are dependent particulars in an analogous way.

3. EXAMPLES OF THE FTB ONTOLOGY

Consider a simple idealized example in electrostatics, that of two isolated point charges q and q' at rest in a vacuum, separated by a distance r. The total electric field is

$$\boldsymbol{E}(\boldsymbol{x}) = \boldsymbol{E}_q(\boldsymbol{x}) + \boldsymbol{E}_{q'}(\boldsymbol{x}). \tag{1}$$

The electric field at point x due to charge q is

$$\boldsymbol{E}_q(\boldsymbol{x}) = \frac{q}{r_x^2}\boldsymbol{e}_x \tag{2}$$

where $\boldsymbol{e}_x$ is the unit vector from q to x and r_x is the distance from q to x. Here is a law of nature concerning the force $\boldsymbol{F}_{q'}$ on test charge q' at point $x_{q'}$

$$\boldsymbol{F}_{q'} = q'\boldsymbol{E}_q(\boldsymbol{x}_{q'}) \tag{3}$$

$\boldsymbol{E}_q$ is composed of a set of $\boldsymbol{E}_q$ kernels. These are $\text{compresent}_{\text{EX}}$ with $\boldsymbol{E}_{q'}$ kernels and with G-x kernels. For simplicity, we consider each kernel non-relationally, that is, as an independently-existing individual.

The electric field $\boldsymbol{E}_q$ produces and explains the force on q', and the FTB ontology provides an account of how it does so. Each kernel contains, among other things, a trope that causes a charge to feel the force described in (3) when the charge is $\text{compresent}_{\text{EX}}$ with the kernel. The independent existence of the $\boldsymbol{E}_q$ and $\boldsymbol{E}_{q'}$ is accounted for in terms of the contingent $\text{compresence}_{\text{EX}}$ of their field kernels. The distinctness of exactly-alike $\boldsymbol{E}_q$ field values is accounted for by the fact that each

field kernel contains an indexing trope. In this way, the FTB ontology for electrostatics accounts for a number of physical features of this example. By contrast, a substance-attribute ontology requires an immaterial substratum to account for these features. In contemporary physics the spacetime manifold is supposed to play the role of immaterial substratum, but, as we saw in Section 1, that ontology faces significant challenges.

When we move to the quantum context, it is plausible that the FTB ontology will enjoy even more significant advantages over substance-attribute ontologies, since this context is quite hostile to traditional notions of substance. Elsewhere I have argued that canonical quantum field theory (QFT) should be understood as a theory about physical fields. I introduced the vacuum expectation value (VEV) interpretation of QFT, on which VEVs for field operators and products of field operators correspond to field values in physical systems (Wayne, 2002). The FTB ontology provides a promising ontology for QFT on the VEV interpretation.

The VEV interpretation of QFT begins by noting that the standard formulation of QFT contains a set of spacetime-indexed field operators for each quantum field. Consider a simple model for quantum field theory consisting of a single non-interacting, neutral scalar quantum field described by a set of spacetime-indexed Hermitian operators $\Phi(\boldsymbol{x}, t)$ that satisfy the Klein–Gordon operator-valued equation. In this model, certain expectation values play a crucial role. These are simply the expectation values for the product of field operators at two distinct points in the vacuum state, $<0|\Phi(\boldsymbol{x}, t)\Phi(\boldsymbol{x}', t')|0>$. These vacuum expectation values (VEVs) describe facts about the unobservable quantum field that have measurable consequences. In particular, one can calculate the probability amplitude of the emission of a quantum of the meson field in a small region around $(\boldsymbol{x}, t)$ and its subsequent absorption in a small region around $(\boldsymbol{x}', t')$ as an integral over appropriate two-point VEVs. This probability amplitude contributes directly to processes which involve the meson field as a mediating force field. It should be noted that it is in fact a Lorentz-invariant combination of two-point vacuum expectation values which plays a role in models of quantum field theories of interest to physicists. A time-ordered product $T\{\Phi(x)\Phi(x')\}$ of field operators can be defined, and the time-ordered two-point VEV $<0|T\{\Phi(x)\Phi(x')\}|0>$ is the covariant Feynman propagator for the meson field, integrals over which are represented graphically by a line in a Feynman diagram. This propagator plays an important role in the derivation of experimentally testable predictions from the model using covariant perturbation theory.

Two-point VEVs provide a perspicuous way to interpret one part of the physical content of our model. On the interpretation being proposed here, these two-point VEVs describe field values in models of physical systems containing quantum fields (although, as we shall see below, two-point VEVs correspond only to a subset of the field values in these models). As mentioned in the previous paragraph, two-point VEVs contribute to probabilities for joint emission and absorption of a quantum of the meson field. They also contribute to probabilities for values of other observables formed as products of field operators, such as total energy and momentum. In this way, field values in the meson model correspond to

physical field values that play the desired ontological role: the field values produce and explain observed subatomic phenomena.

The central claim of the VEV interpretation of quantum field theory holds that VEVs in standard quantum field theory correspond to field values in physical systems containing quantum fields. It is a useful fact about quantum field theory that certain VEVs offer an equivalent description of *all* information contained in the quantum field operators, their equations of motion and commutation relations. In general, a set of VEVs uniquely specifies a particular $\Phi(x,t)$ (satisfying specific equations of motion and commutation relations) and vice versa. As Arthur Wightman first showed, for expectation values fully to describe a quantum field operator, one must specify not only VEVs at each point,

$$<0|\Phi(x,t)|0> \quad \text{for all } x,t, \tag{4}$$

but also vacuum expectation values for the products of field operators at two different points,

$$<0|\Phi(x_1,t_1)\Phi(x_2,t_2)|0> \quad \text{for all } x_1,t_1,x_2,t_2, \tag{5}$$

at three points, and so on (Wightman, 1956; cf. Schweber, 1961, 721–742). In these expressions for vacuum expectation values I let $\Phi(x,t)$ stand for the adjoint field as well, $\Phi^{\dagger}(x,t)$; in general, vacuum expectation values contain both field operators and their adjoints. Wightman determined that a complete specification of an interacting quantum field operator requires vacuum expectation values of all finite orders. In Wightman's reformulation of quantum field theory, operator-valued field equations are replaced by an infinitely large collection of number-valued functions constraining relations between expectation values at different spacetime points.

The VEV interpretation highlights three ways in which quantum fields differ from classical fields, and all three of these differences are well accommodated within the FTB ontology. First, VEVs determine probabilities for field values, and these probabilities may be understood as propensities, unlike the classical case in which field values are all definite-valued. This widening of the notion of field value, from a definite value in the classical case to a set of propensities in the quantum case, is naturally accommodated within the FTB ontology. As we have seen, tropes, which are simply particular property or relation instances, include dispositional and propensity instances as well. More precisely, a quantum field is composed of a set of kernels, where each kernel is made up of one or more geometrical G tropes, an indexing x trope, and an F trope, all $\text{compresent}_{\text{IN}}$. Each F trope in a quantum field is a propensity for an n-point field value, and each F trope corresponds, in the VEV interpretation, to one n-point vacuum expectation value. Each kernel is itself $\text{compresent}_{\text{EX}}$ with other kernels making up the quantum field, corresponding to other n-point values, and with kernels of independent fields. As we have seen, $\text{compresence}_{\text{EX}}$ is an external relation constituted by one or more relational E tropes.

Secondly, quantum fields contain single-point and n-point field values, understood as n-point kernels on the FTB approach. This is in contrast with classical fields, which consist exclusively of single-point values. Thus F tropes in quantum

fields are irreducibly relational in a way that F tropes in classical fields are not (recall that F tropes in classical fields may be relational in another way, namely in their dependence on the neighbourhood of a point). As noted above, the FTB approach naturally accommodates this expansion of the kernel to regions in order to include these compresent$_{\mathrm{IN}}$ relational tropes. Indeed, because quantum fields contain n-point VEVs of all orders there may be no finite region of a quantum field that is separable from the rest of the field. This is accommodated by such a quantum field having a single kernel.

Thirdly, a quantum field does not determine the state of the system. The actual state of a physical system containing a quantum field corresponds to a specific state vector/operator combination, yet on the VEV interpretation the state vector plays no role in specifying the field values of a quantum field. The implication for the FTB ontology is that the kernel of a quantum field cannot exist independently of some additional tropes, those comprising the state of the system. A quantum field kernel must be compresent$_{\mathrm{EX}}$ with state tropes, and the kernel depends for its existence on compresence with some trope of the state-trope type.

4. CONCLUSION

Ontological parsimony, flexibility, and moderate nominalism are attractive features of field trope-bundle ontologies for field theories in physics. FTB approaches have significant advantages over traditional substance-attribute approaches, and this chapter has sketched, in a very preliminary way, how such trope-bundle ontologies can be constructed for classical and quantum field theories. Clearly, much work remains to be done to flesh out these constructions. However, I hope to have shown that these ontologies are promising choices to underwrite the causal and explanatory roles physical fields play in contemporary physics.

ACKNOWLEDGEMENT

I would like to thank Michal Arciszewski, Gordon Fleming, Storrs McCall, Ioan Muntean, Paul Teller and the audience at the Second International Conference on the Ontology of Spacetime for helpful discussion and comments on earlier drafts of this chapter.

REFERENCES

Belot, G., 2000. Geometry and motion. The British Journal for the Philosophy of Science 51 (Supp), 561–595.

Campbell, K., 1990. Abstract Particulars. Blackwell, Oxford.

Daly, C., 1994. Tropes. Proceedings of the Aristotelian Society 94, 253–261.

Earman, J., 1989. World Enough and Space-Time: Absolute Versus Relational Theories of Space and Time. Cambridge, Mass, MIT Press.

Earman, J., Norton, J., 1987. What price spacetime substantivalism: The hole story. British Journal for the Philosophy of Science 38, 515–525.

Field, H., 1980. Science without Numbers. Princeton University Press, Princeton.
Field, H., 1989. Realism, Mathematics, and Modality. Blackwell, Oxford, UK.
Green, G., 1838. On the laws of the reflexion and refraction of light at the common surface of two non-crystallized media. Transactions of the Cambridge Philosophical Society 7 (1), 113.
Kelvin, W.T., 1904. Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light. C.J. Clay and Sons, London, Baltimore.
Larmor, J., 1900. Aether and Matter. Cambridge University Press, Cambridge.
Lorentz, H.A., 1895. Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern. E.J. Brill, Leiden.
Lorentz, H.A., 1909. The Theory of Electrons. B.G. Teubner, Leipzig.
Malament, D., 1982. Review essay: Science without Numbers by Hartry Field. Journal of Philosophy 79, 523–534.
Molnar, G., Mumford, S.E. (Eds.), 2003. Powers: A Study in Metaphysics. Oxford University Press, Oxford.
Moreland, J.P., 1997. A critique of Campbell's refurbished nominalism. Southern Journal of Philosophy 35, 225–245.
Parsons, G., McGivern, P., 2001. Can the bundle theory save substantivalism from the hole argument? Philosophy of Science 68 (3), S358–S370.
Redhead, M., Teller, P., 1991. Particles, particle labels, and quanta: The toll of unacknowledged metaphysics. Foundations of Physics 21, 43–62.
Rynasiewicz, R., 1996. Absolute versus relational spacetime: An outmoded debate? Journal of Philosophy 93, 279–306.
Schaffner, K.F., 1972. Nineteenth-Century Aether Theories. Pergamon Press, Oxford, New York.
Schweber, S.S., 1961. An Introduction to Relativistic Quantum Field Theory. Row Peterson, Evanston, IL.
Simons, P., 1994. Particulars in particular clothing: Three trope theories of substance. Philosophy and Phenomenological Research 54, 553–575.
Teller, P., 1991. Substance, relations, and arguments about the nature of spacetime. The Philosophical Review 100 (3), 363–396.
Teller, P., 1995. An Interpretive Introduction to Quantum Field Theory. Princeton University Press, Princeton.
Wayne, A., 2002. A naive view of the quantum field. In: Kuhlmann, M., Lyre, H., Wayne, A. (Eds.), Ontological Aspects of Quantum Field Theory. World Scientific, Singapore, pp. 127–133.
Wightman, A.S., 1956. Quantum field theory in terms of vacuum expectation values. Physical Review 101, 860–866.
Williams, D.C., 1997. On the elements of being: I. In: Mellor, D.H., Oliver, A. (Eds.), Properties. Oxford University Press, Oxford, New York, pp. 112–124.

CHAPTER 2

Is Structural Spacetime Realism Relationism in Disguise? The Supererogatory Nature of the Substantivalism/Relationism Debate

Mauro Dorato*

Abstract

In this chapter I position the substantivalism/relationism debate in the wider context of the scientific realism issue, and investigate the place of structural realism in this debate.

This chapter tries to connect the substantivalism/relationism debate to the wider question of scientific realism. Historically, the issue of the reality of spacetime (substantivalism) was certainly fuelled by a more favourable attitude toward scientific realism, which emerged after the crisis of the neopositivistic criterion of meaning during the second half of the 20th century.

However, there are not just historical reasons for exploring the above connection in a more systematic way. On the one hand, within the camp of scientific realism, in the last couple of decades *structural realism* has emerged as a sort of *tertium quid* between a radically sceptical antirealism about science and an allegedly "naïve realism" about the existence of theoretical entities.[1] On the other, difficulties to adjust the substantivalism/relationism dichotomy to the framework of the General Theory of Relativity (GTR) have pushed philosophers of space and time to find alternative formulations of the debate. Among these, various forms of *structural spacetime realism*—more or less explicitly formulated—have been proposed either as a third stance between the two age-old positions (Stachel, 2002; Rickles and French, 2006; Esfeld and Lam, 2006), or as an effective way to overcome or dissolve

* Department of Philosophy, University of Rome 3, Italy

[1] In Worrall's original view (1989), for example, structural realism was meant to give an account of both the predictive success of science and of its continuity across scientific change, while granting Laudan's pessimistic meta-induction against the existence of theoretical entities (Laudan, 1981).

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00002-8
© Elsevier BV
All rights reserved

the substantivalism/relationism debate (Stein, 1967; DiSalle, 1995; Dorato, 2000; Dorato and Pauri, 2006; Slowik, 2006).

The attempt at using structural realism in order to steer a middle course between substantivalism and relationism and to defend a structural form of realism about spacetime, however, raises several questions.[2] One of these is the following: if structural realism claims that "science is about structure", or about physical *relations* that are partially described by our mathematical models of the physical world, in what sense is structural spacetime realism really different from good old relationism?

My main answer to this crucial question will be two-fold:

(1) Viewed from the perspective of the substantivalism/relationism debate, structural spacetime realism (i.e., the claim that spacetime is exemplified structure) is a form of relationism;
(2) However, if we managed to reinforce Rynasiewicz's (1996) point that GTR makes the substantivalism/relationism dispute "outdated", the re-elaboration of Stein's 1967 version of structural spacetime realism to be proposed here proves to be a good, *antimetaphysical* solution to the problem of the ontological status of spacetime.

In short, it is only if we assume that the dispute between substantivalism and relationism is still meaningful also in the context of GTR that structural spacetime realism turns into a form of relationism. But since that dispute will be shown to be unfit for GTR, structural spacetime realism gives a good answer to the problem of the status of spacetime that is neither relationist nor substantivalist, and overcomes both positions.

The chapter is divided into three parts. In the first (Section 1), I briefly review the main positions in the game of scientific realism, with the intent of showing that *if* the substantivalism/relationism is genuine, then structural spacetime realism is a form of relationism (*first claim*). In the second part (Section 2), I reconstruct what I take to be Stein's (1967) position on the ontological status of spacetime and on the related issue of scientific realism. While he in no way was explicitly trying to defend structural spacetime realism as it is now discussed, I will argue that, especially after the onset of GTR, Stein's claim that worrying about the ontological status of the exemplified structure is "supererogatory" (superfluous or otiose) proves quite robust against four foreseeable objections.

Finally, in Section 3, I will show how the duality of the metric field and the difficulties of defending a "container/contained", or a "spacetime/physical field" distinction in classical GTR speak definitely in favour of a dissolution of the substantivalism/relationism debate, and therefore of a structural realist solution to the question of the ontological status of spacetime (*second claim*).

[2] For some of these, see Pooley (2006).

1. THREE FORMS OF SCIENTIFIC REALISM AND THEIR CONCEPTUAL RELATIONSHIPS

Schematically, there are three versions of scientific realism in the current philosophical debate, whose logical and conceptual relationships are the target of ongoing controversies. In this section, I will briefly sketch the three positions, by dedicating somewhat more attention to the tenets of structural realism. This will prove necessary to situate this doctrine in a wider conceptual framework, and thereby gain a deeper understanding of its main implications.

(1) According to *theory realism*, well-confirmed theories are *true*, either *tout court*, or *approximately*, i.e., in the approximation of the model.

The crucial term in this position is obviously "approximately true": if one decides to forgo as being too audacious the claim that theories are true "without qualifications", one encounters various problems in giving a precise account of the notion "approximate truth" (see, for instance, Niiniluoto, 1999, Section 3.5; Smith, 1998, Chapter 5; Psillos, 1999 Chapter 11).[3] Given my purposes, I will simply leave these difficult questions by side, and move on to the second form of scientific realism.

(2) *Entity realism*: "theoretical", *non-directly* observable entities postulated by well-confirmed theories (quarks, muons, electrons, black holes, etc.) have a mind-independent existence.

As is evident, this definition presupposes a distinction between what is observable with the naked eye and what is observable only with the help of instruments. *Entity realists* typically note that electrons *are* observable, albeit indirectly. If the distinction between direct and indirect observability is one of degree and therefore not ontologically significant, in their opinion we should believe in the existence of electrons or quarks for the same reasons that we grant mind-independent existence to tables and chairs: not only do we perceive them (although indirectly), but we measure and manipulate them to obtain our aims. Antirealists about entities typically use evidence from past science to draw our attention to the numerous entities that have been abandoned during its history (flogist, caloric, aether, etc.). They then note that the methodology used by past theories that postulated what we now regard as non-existing entities is the same that we used today. Consequently, according to the entity antirealist, we should abstain from believing in the theoretical components of current physical models, but only accept them as being *empirically adequate* (Van Fraassen, 1980).

(3) Structural realism claims that science is about structures: while structures are real and knowable, entities—*if regarded as endowed only with monadic properties*—are either unknowable or unreal.

Structural realists have not been always very clear about the nature of *physical versus* purely *mathematical* structures. Following Poincaré, in this chapter I will

[3] Supposing with Popper that we don't know whether our current theories are true, how can we estimate their distance from the true theories? Furthermore, does the notion of "being truth" (or "being false") admit of degrees?

understand the former as a class of *physical relations* partially described by the latter, that is, by the equations or laws defining a mathematical model:

> «**The differential equations are always true**, they may be always integrated by the same methods, and the results of this integration still preserve their value. It cannot be said that this is reducing physical theories to simple practical recipes; **these equations express relations, and if the equations remain true, it is because the relations preserve their reality**. **They teach us now, as they did then, that there is such and such a relation between this thing and that**; only, the something which we then called *motion*, we now call *electric current*. But these are merely names of the images we substituted for the real objects which Nature will hide for ever from our eyes. **The true relations between these real objects are the only reality we can attain**, and the sole condition is that the same relations shall exist between these objects as between the images we are forced to put in their place. If the relations are known to us, what does it matter if we think it convenient to replace one image by another?» (Poincaré, 1905, pp. 160–1, the emphasis in bold is mine)

Note that Poincaré does not deny the existence of "real objects" or theoretical entities; rather, he simply declares them to be unknowable ("the real objects which Nature will hide for ever from our eyes"). Consequently, following Ladyman, we can distinguish two forms of structural realism: depending on whether the concrete, physical relations partially referred to by mathematical models are the only things we can know (Poincaré, 1905; Worrall, 1989), or are regarded as the only existing stuff (French and Ladyman, 2003; Esfeld, 2007; Esfeld and Lam, 2006), we have *epistemic* or *ontic* structural realism (Ladyman, 1998).

In the former, *epistemic* case, entity realism is not denied, but possibly reached at "the limit of inquiry", as more and more relations between objects are discovered (Cao, 2003). Epistemic structural realism can therefore be either *agnostic* about theoretical entities, or simply presuppose them, with Poincaré, as the indispensable but unknowable *relata* of the relations described by and known *via* scientific theories and laws.

In the *ontic* version of structural realism, instead, entity realism is simply outlawed: entities, *if* regarded as bearers or bundles of, monadic, intrinsic properties, are "crutches" to be thrown away after the construction of the model. Ontic structural realism, as I understand it, is a form of *atheism* about entities, but only if the latter are conceived as endowed with intrinsic, monadic properties in the sense of Langton and Lewis (1998). Roughly speaking, an intrinsic, monadic property is a property that, like boldness, can be attributed to an individual without presupposing the existence of another individual. A property is extrinsic or relational if and only if it is not intrinsic.

This interpretation of ontic structural realism seems to be shared by structural realists like Esfeld (2007) and Esfeld and Lam (2006): since *ontic* structural realists cannot be radically instrumentalist about the referential import of models, they must redescribe all ontological claims of moderns science in such a way that theoretical entities simply turn out to be *bundles of relations*. In this version of the

theory, the *relata* of the relations described by science are bundles of relations, and it is therefore accepted that relations cannot exist without their *relata*. In this way, one of the standard objections raised against a more radical view of ontic structural realism (French and Ladyman, 2003) is tackled. However, it seems to me that it is possible to read also French and Ladyman as defending *this* version of ontic structural realism, since even the bundles of relations on which the radical ontic structural realists bets are, after all, *entities* of some kind.[4] I daresay that no ontic structural realist should be falling into the trap of accepting the view that "relations can exist without relata".[5]

Epistemic structural realism has its problems: one may legitimately wonder with Esfeld and Lam (2006) whether it is reasonable to detach epistemology from ontology in such a radical way as to postulate entities that—similar to kantian noumena—are endowed with intrinsic properties that in principle we will never know. But ontic structural realism, even in the moderate form postulated by Esfeld and Lam (2006), is not without troubles, as it is natural to raise doubts about whether an ontology of "entities" possessing purely relational properties is plausible.

For example, one might question whether entities can bear relations to one another without having any intrinsic properties whatsoever: the relation "*a* is heavier than *b*" presumably holds because of a property like "having a certain density of matter", that seems intrinsic to each and every body.

However, independently of conceptual difficulties of this kind, the main point of structural realism in both versions is that their defenders agree that it is natural science that should decide in favour or against the epistemic inaccessibility or the non-existence of intrinsic properties, and not just armchair, *a priori* conceptual analysis.

For instance, if *mass*, *spin* and *charge* could be legitimately regarded as *intrinsic* properties of elementary particles, ontic structural realism as I presented it would be automatically refuted. *Prima facie*, it is hard to see why these should not qualify as *bona fide intrinsic* property of particles, even though, of course, to get to know them, we must have other entities interact with them. Analogously, if we granted that these three properties, treated as causal powers of the entities possessing them, are reliably *known* by current physical theories, also epistemic structural realism would be rejected: we could know at least *some* intrinsic properties of some theoretical entities.

Also the case of entangled particles, considered by ontic structural realism as paramount evidence for their position (Esfeld, 2004), should be discussed in light of a *dispositionalist* interpretation of quantum mechanics. If the quantum properties of entangled particles could be regarded as *dispositional*, then even moderate ontic structural realism should be re-evaluated, since such dispositional properties, belonging to any quantum entity in a superposed, entangled state, should, *pace* Popper, be regarded as *intrinsically* possessed (Dorato, 2006b; Suárez, 2004).

[4] The distinction between radical and moderate ontic structural realism is in Esfeld and Lam (2006).

[5] For this view, a form of which could perhaps be attributed to David Mermin, see Barrett (1999, p. 217).

In a word—except for some inevitable vagueness in the distinction between intrinsic and extrinsic properties, which blurs the distinction between entity realism and ontic structural realism—the requisites of structural realism are sufficiently strict. Unfortunately, a thorough study of structural realism *vis à vis* the properties of particles within the standard model is yet to be written. The same conclusion holds for the consequences of a dispositionalist interpretation of quantum mechanics on the relational ontology of structural realism.

However, even if the confrontation with field theory were to result in a negative verdict, one could still imagine that *some* version of structural realism could survive if applied to spacetime physics.

This is exactly the issue that I will try to explore in the remainder of this chapter: "local" philosophical analysis may sometimes be more interesting than sweeping and vague attempts at encapsulating the whole of science or of physics in one scheme. Structural realism may fail as metaphysics for quantum field theory and yet be successful for spacetime physics: if this were the case, we would simply have another piece of evidence in favour of the metaphysical disunity of science. After all, it would be strange to find out that a single metaphysical claim squared with both quantum theory and spacetime physics, given that these two theories have not yet been reconciled in a single frame.

Before passing to the definitions of substantivalism, let me briefly note the logical relationships of these various forms of scientific realism. Quite naturally, a theory cannot even be approximately true if the entities and the structure it postulates don't exist at all. This shows that theory realism implies both entity realism and structural realism, so that 1) implies 2) and 3). Since, by contraposition, ¬2) implies ¬1), if 3) implied ¬2), 3) should also deny 1). Now, since ontic structural realism ought to be regarded as a denial of the existence of entities endowed with intrinsic properties (entity antirealism), it also entails theory antirealism, given that we just showed that ¬(2) implies ¬(1).

However, if the only existing *entities* were bundles of relations, ontic structural realism would trivially degenerate into entity realism and would be trivially compatible with it. Epistemic structural realism, on the other hand, is definitely not against the reality of the *relata*, but simply insists on their epistemic accessibility. As a consequence, structural realism in its various forms is compatible with entity realism, but not committed to theory realism, at least to the extent that entity realism, as some philosophers have it, is compatible with instrumentalism about theories and laws.

1.1 Substantivalism and structural spacetime realism

In order to understand the implications of the two forms of structural realism for the nature of spacetime, we need precise definitions of both "substantivalism" and "substance". In the literature on GTR, we find two main types of substantivalism, "manifold substantivalism" and "metric field substantivalism", depending on whether spacetime is identified with the differentiable manifold or with the metric field (plus the manifold):

> MANIFOLD SUBSTANTIVALISM «Space-time is a substance in that it forms a substratum that underlies physical events and processes, and spatiotemporal relations among such events and processes are parasitic on the spatiotemporal relations inherent in the substratum of spacetime points and regions.» (Earman, 1989, p. 11)
>
> METRIC FIELD SUBSTANTIVALISM «A modern day substantivalist thinks that space-time is a kind of thing which *can*, in consistency with the laws of nature, exist independently of material things (ordinary matter, light and so on) and which is properly described as having its own properties, over and above the properties of any material things that may occupy parts of it.» (Hoefer, 1996, p. 5, my italics)

Relationism is a denial of these two theses, and if both definitions of spacetime substantivalism were legitimate, it would come in two forms. While relationism about the manifold would be consistent with metric field substantivalism, a denial of the latter view would seem to entail also a denial of manifold substantivalism.

Note that, in the first definition, spacetime is a substance in virtue of its being a *substratum* underlying physical events, a position which certainly refers to one of the traditional meanings of "substance".[6] The second definition seems to presuppose a second sense of "substance", as something existing independently of other entities and events.[7]

Manifold substantivalism is based on the presupposition that the very debate between substantivalism and relationism requires a clear-cut *separation between spacetime*—regarded as a *container—and physical systems*, gravitational and non-gravitational ones alike, regarded as whatever is contained in it. As we will stress in Section 3, and as noted already by Rynasiewicz (1996), this definition of substantivalism creates conceptual troubles to the extent that GTR "overcomes"[8] the separation between container and contained for reasons that will become clear in Section 3.

The second definition, capturing *metric field substantivalism*, relies on Einstein's field equation, which allows us to write the gravitational field and ordinary matter on the two different sides of the equation. The italicized "can" of the second quotation refers to the fact that *the metric field can exist without matter*, even though it is typically correlated with it by Einstein's equations. This second definition creates controversies to the extent that it identifies *spacetime* with the manifold and the metric field, the metric field in GTR being a *physical* field, that one might want to regard (erroneously, in my view) as something being "contained" in something else enjoying an independent existence (the manifold).

Equipped with these definitions of scientific realism and substantivalism, we are now ready to try to understand the consequences of structural realism as applied to spacetime (i.e., structural spacetime realism) *vis à vis* the substantivalism/relationism debate, *assuming*, for the time being, *that such a debate is genuine*.

6 From the Latin *sub stare,* to lie under.

7 On this second sense of substance, more below.

8 "Overcome" here corresponds to the technical sense rendered by the German verb *aufheben* in Hegel's philosophy: it is an overcoming that somehow realizes a synthesis of the views that were previously regarded as opposed and irreconcilable.

According to an *epistemic* version of structural spacetime realism, spatiotemporal relations would be all that can be known about spacetime: the nature of the *relata* (points, physical events), together with their first order, intrinsic properties, would be unknowable (as Poincaré had it, they would be "for ever hidden from our eyes"). In the *ontic* version of structural spacetime realism, spatiotemporal relations would instead be *all that there is*: spacetime points or physical events endowed with intrinsic properties would simply not exist, and would have to be re-conceptualised in terms of relations. From this perspective, a point P would just *be* something bearing the spatiotemporal relations $R_1, R_2, \ldots, R_n$ to other n points, and these relations would constitute its identity.

I will now argue that—independently of whether spacetime is represented by the manifold or by the manifold plus the metric field—if we think that the dispute between substantivalists and relationists is genuine also after GTR, *structural spacetime realism is a form of relationism*.

Prima facie, this conclusion seems less justified for epistemic structural spacetime realism (let me use the acronym ESSR). It will be recalled that it claims that spatiotemporal points might, or even should, exist *qua relata* of the spatiotemporal relations, but that we will never get to know their intrinsic properties: it is only their spatiotemporal relations that are epistemically accessible. To the extent that substantivalism implies the existence of spatiotemporal points endowed with *intrinsic* properties, ESSR could coherently defend it, but would have to consider it as a metaphysical doctrine which could be never confirmed or disconfirmed by empirical science.

As a consequence of the fact that the defenders of ESSR must leave substantivalism beyond the reach of empirical science, they seem to be facing a choice between two alternatives. The first consists in dropping the substantivalist/relationist debate altogether as irrelevant for empirical science, which leads us very close to the second claim to be argued for in the following (Section 3). The second alternative consists in embracing ontic structural spacetime realism, i.e., move toward a position that brings a structuralist epistemology into line with a metaphysics postulating *just* the existence of relations.

In a word, ESSR *per se* is certainly compatible with substantivalism, but looks like a remarkably unstable philosophical position. If one does not drop the dispute (first alternative), or does not opt in favour of ontic structural spacetime realism (second alternative), the compatibility with substantivalism would be purchased at too high a price, as it would amount to buying a metaphysical theory that could not be measured *in principle* against the results of a physical theory.

I will now show how also *ontic* structural spacetime realism (call it OSSR, the second alternative mentioned above), with its denial of the existence of intrinsic properties, is against the existence of a substantival spacetime, and turns into pure relationism. Since my argument crucially hinges on the assumption that by "substance" we should mean an entity endowed with intrinsic properties, *i.e.*, something that exists independently of any other entity, we must ensure that this definition is reasonable also in the context of spacetime physics. In order to do so, two remarks are appropriate.

The first remark is that the philosophical tradition yields a univocal verdict with respect to the meaning of "substance": the main difference between an *accident* like being married and a *substance* like Socrates is that the latter, unlike the former, exists independently of anything else. Descartes—to name just one of the philosophers who played an essential role in transplanting the Aristotelian tradition into the soil of modern philosophy—tells us that "when we conceive a substance, we understand nothing else than an entity which *is* in such a way that it needs no other entity in order to *be*." (Descartes, 1644, I). A very similar definition of substance has been defended also by Spinoza: "Per substantiam intelligo id quod in se est et per se concipitur…" (Spinoza, 2000, I, Prop. 3)[9]. And these are but two examples.

If we accept this definition of substance, we should attribute a substantial spacetime (or a region of it, up to a single point) *intrinsic* properties, i.e., properties that can be attributed without presupposing the existence of other entities. This would be sufficient to show that ontic structural spacetime realism is incompatible with point-substantivalism, and is a form of relationism.[10]

The same result is derivable if we use "substance" to refer to an entity possessing a distinct identity, or an individuality derived by the possession of some intrinsic property. In this second, closely related sense of "substance", spatiotemporal points are substantial if and only if they have a distinct identity *just taken by themselves*.

Relative to this second sense of substance, Stein (1967) has first shown how both Leibniz *and* Newton denied substantiality to points and instants: also according to Newton, points and instants receive their identity from the spatiotemporal order to which they belong, as each is qualitatively identical to any other.[11] It follows that any ontology denying the existence of intrinsically individuating, monadic properties is anti-substantival or relational also in the context of spacetime physics. But according to the ontic structural spacetime realist, a point or an instant has no other individuality than that of being in relation to other points: taken by itself, it has no identity and is therefore not substantial also in this second sense of substance.

The above mentioned second remark conceives the possibility of a different definition of "substance", one that would justify the neutrality of the substantivalism/relationism debate with respect to structural spacetime realism. After all, one could argue, in changing scientific contexts it is unavoidable that even notions with an important historical tradition be readjusted to fit a new conceptual framework. However, words have meanings, and contrary to the opinion expressed by Humpty Dumpty in *Alice in wonderland*, we cannot have them mean what *we* want.

And even if in the present case such a change could be done, the dispute about the substantial or relational nature of spacetime would be transformed into

[9] By substance I mean something which exists by itself and can be conceived by itself…, my translation).

[10] This remark counters an objection raised by Michael Esfeld in his reading of a previous version of this chapter.

[11] In the unpublished manuscript *De Gravitatione et equipondio fluidorum*, Newton writes: "the parts of space derive their character from their positions, so that if any two could change their positions, they would change their character at the same time and each would be converted numerically into the other qua individuals. The parts of duration and space are only understood to be the same as they really are because of their mutual order and positions (*propter solum ordinem et positiones inter se*); nor do they have any other principle of individuation besides this order and position which consequently cannot be altered" (Janiak, 2004, p. 25).

a purely *semantic* question, depending on the meaning of "substance". I think it is fair to add that when a philosophical question turns into an issue pertaining exclusively to the meaning of words, then it tends to be deprived of much of its significance. Since this is the second claim that I want to defend in my chapter, I will postpone its defence in Section 3: for the time being it is sufficient to have illustrated the point that it is difficult to escape from the traditional meaning of "substance" as something that exists independently by possessing intrinsic properties. This fact pushes OSSR in the arms of relationism.

Despite these remarks, my fist claim (that structural spacetime realism is a form of relationism) might have been established too quickly. The well-known independence/autonomy of the metric field from the matter field might seem to speak against my view, since in *empty* solutions to Einstein's field equations, the metric field would seem to enjoy the status of an independently existing *substance* (metric field substantivalism). Could the moderate form of OSSR defended by Esfeld and Lam (2006) be compatible with metric field substantivalism, or even be *neutral* with respect to the substantivalism/relationism dispute?

Let us recall that according to OSSR, the entities exemplifying the spatiotemporal/metric relations do *not* possess intrinsic properties (or a primitive thisness) over and above that of standing in certain spatiotemporal relations. That is, these entities are nothing but that which stands in these relations. Against my view, it could then be argued that OSSR need not take a stand about the question of what these entities are: *they might be spacetime points (substantivalism) or material entities, namely parts of the matter fields (relationism)*.[12]

I already granted that the gravitational field and the non-gravitation field can have a distinct existence, since, in $T = 0$ solutions, the former field *can* exist without the latter. According to the previous approach to the notion of substance, if we consider *the whole* of the metric field, shouldn't we regard it as a *substance*, even if its *parts* (points), as the ontic structural spacetime realist has it, do not possess intrinsic properties, or independent existence?[13]

This question can be tackled in at least four different ways:

(i) OSSR cannot be made compatible with metric field subtantivalism. In fact, if the metric field as a *whole* exists as a substance and has therefore an independent existence, presumably it would have intrinsic properties, as all substances have, namely properties attributable to the metric field as a whole independently of anything else. But such an "intrinsicness" or independence of the metric field from the matter field would be hardly compatible with OSSR's relationalist ontology. An entity without relations to something else can hardly be admitted within the latter ontology. And my first claim would be vindicated.

(ii) Suppose the metric field as a whole is substantial while its parts aren't. How can the whole of the metric field be a substance if its parts (regions and points) cannot have intrinsic properties in virtue of the requirements of a structuralist ontology? Typically, the parts of compound substances are themselves

[12] This way of putting the issue was suggested by Michael Esfeld in his comments.

[13] For this holism of the metric field, see Lusanna and Pauri (2006).

substances: the pages of a book (a composite substance) are themselves substances, whether they are detached from the book or not; but if spacetime substantivalism required the existence of spacetime points or region as individual substances, it would go against ontic structural spacetime realism! The dilemma in which the defender of OSSR inclined toward substantivalism is caught is not easily solvable: if the whole metric field *is* a substance, then it must have intrinsic properties. But then the compatibility with OSSR is lost. On the other hand, if the whole metric field is *not* a substance, OSSR has a living chance, but its compatibility with substantivalism is lost, because (contrary to what is actually the case), the metric field would not be independent of the matter field. In either horns of the dilemma, my first claim is vindicated.

(iii) If the substantivalist/relationist debate were simply a matter of deciding whether the gravitational field is distinct and independent from the matter field or not, relationalism couldn't win, and GTR would be substantivalist by *fiat*, without even beginning to fight. This way of cashing the debate would trivialize it. Of course, from the fact that it has such an easy solution, we cannot conclude that the debate is outdated. However, since the question whether GTR is substantivalist, relationist or neither will be evaluated in Section 3, we can move to the last reply, which addresses Esfeld's proposed alternative between the relata of the spatiotemporal relations being *spacetime points* (substantivalism) or *parts of the matter fields* (relationism).

(iv) The expression "spacetime points" is not unambiguous, as it has at least two distinct interpretations. If by "spacetime points" one meant points of the manifold endowed with primitive identity or *intrinsic* properties, one would have *manifold substantivalism*. Since this position would contradict ontic structural spacetime realism, it cannot be the intended interpretation. On the other hand, if by "spacetime points" one meant points of the metric field, one would have to decide whether such a field is geometrical/spatiotemporal or physical (i.e., substantival or relational). Since, as we are about to see in Section 3, the main lesson of GTR is that it is *both*, it is hard to make sense of the question whether we have a substantival spacetime (because spatiotemporal points *exist* on their own as individuated by their metric relations) or a relational spacetime (because spatiotemporal relations supervene on the gravitational field, which is a *physical* field).[14]

Since the question of the status of metric field vis à vis spacetime will be discussed below, here I can afford ending my discussion with two quotations, which illustrate the connection between structural spacetime realism and relationism in a particularly clear way: "There is no such thing as an empty space, *i.e.* a space without field. Space-time does not claim existence on its own, but only as a structural quality of the field" (Einstein, 1961, pp. 155–156); "spacetime geometry is nothing but the manifestation of the gravitational field" (Rovelli, 1997, pp. 183–184). Despite the fact that argument from authority have no value even if they come from

[14] The above ambiguity is not present in making sense of the claim that "metric relations amount to relations among material entities" (relationism, as Esfeld has it), since "material points" should mean "points of the matter-field". In this "leibnizian" interpretation, however, one is forbidding pure gravitational solutions to the Einstein's field equations; in view of the existence of $T = 0$ solutions of such equations, this seems too high a price to pay to have a plausible formulation of the substantivalism/relationism debate also in the context of GTR.

Einstein, it should be admitted that as expressions of structural spacetime realism, these quotations also look like acts of relationist faith!

In conclusion, structural spacetime realism either pushes toward, or just *is, relationism*, and in any case it cannot it be regarded as a *tertium quid* between substantivalism and relationism.

However, if we were to agree that the substantivalism/relationism dichotomy has no clear-cut application within GTR, we would need an alternative formulation of the problem of the nature of spacetime, more attentive to the *ontological* problem of its existence than to the *metaphysical* question of a substantival *vs.* a relational existence. As we are about to see in the next section, the seeds of such an important but neglected anti-metaphysical formulation are to be found in Stein (1967), and need to be developed and defended against possible criticism.

2. A REFORMULATION OF THE SUBSTANTIVALISM/RELATIONISM DEBATE: STEIN'S VERSION OF "STRUCTURAL SPACETIME REALISM"

> «If the distinction between inertial frames and those that are not inertial is a distinction that has a real application to the world; that is, if the structure I described[15] is in some sense really exhibited by the world of events; and if this structure can legitimately be regarded as an explication of Newton's "absolute space and time"; then the question whether, in addition to characterizing the world in just the indicated sense, this structure of space-time also "really exists", surely seems supererogatory» (Stein, 1967, p. 193)

Let us recall that a supererogatory (*überverdienstlich*) action, according to the *Critique of Practical Reason* for Kant is an action that goes beyond what is required by one's duty, despite its being possibly inspired by noble sentiments. In a word, according to Stein, worrying about the independent existence of the exemplified structure is otiose. This is the position I would like to defend.

By using a later paper of his (Stein, 1989), I read Stein as claiming that the traditional dispute between substantivalism and relationism is analogous to that between scientific realism and antirealism as he viewed it: *neither position is tenable*! If antirealism about spacetime structure amounted to a position denying that the world of events "really exhibits" a certain geometrical or spatio-temporal structure, something that Stein instead explicitly grants, such antirealism about spacetime would *not be* tenable. «The notion of structure of spacetime" is not to be regarded "as a mere conceptual tool to be used from time to time as convenience dictates… there is only one physical world; and if it has the postulated structure, the structure is—by hypothesis—there, once and for all» (Stein, 1967, p. 52).

However, Stein is not "realist" about spacetime either: if spacetime *realism* were equivalent to the supererogatory claim that the spatiotemporal structure "really exist"—where "really exists" presumably refers to the *independent existence* of the structure (over and above the physical events instantiating it) required by

[15] If N denotes the mathematical model for absolute space and time, $N = <S \times T>$, i.e., N is the Cartesian product between the three-dimensional Euclidean space S and the time T.

some forms of substantivalism—such an (hyper-)realism about spacetime structure would not be reasonable either.

Does Stein's position amounts to proposing a *tertium quid* between substantivalism and relationism?[16] I want to push the point that if Stein is right in insisting that the opposition between substantivalism and relationism is *not* a fruitful way to make sense of the Newton–Leibniz debate, and I think he is correct about this, *a fortiori* it is not fruitful within GTR, where there is no empty, container space in the sense presupposed by the ancient atomists. Following Stein's "style" of philosophical analysis as I understand it, I think that the important questions to be raised are:

- What did the "relationist" Leibniz and the "substantivalist" Newton *agree upon?* (according to both, for instance, instants and points have *no* intrinsic identity)
- How do our spatiotemporal models represent the physical world?
- What does it mean to claim that spacetime *exists*?

Since I cannot pursue the first question here, let me expand on the other two, starting from the last. If we agree in stipulating that **"spacetime exists iff the physical world exhibits the corresponding spatiotemporal structure"**, I would like to press the point that the empirical success of our spacetime models do raise an important *ontological* question ("does spacetime exist?"), while the particular manner of existence of spacetime, namely whether it is substance-like or relation-like, after the establishment of GTR has become a less central, *metaphysical, possibly merely verbal* question. I am here relying on a much neglected distinction between ontology and metaphysics: the former addresses question of existence ("what there is"), the latter is involved in the particular manner of existence.[17] A one-sentence way of putting the main point of this chapter would be the following: spacetime exists as exemplified structure, while the question whether it exists as substance or relation is not well-posed.

2.1 Some foreseeable objections to Stein

Once we accept the view that spacetime structure postulated in mathematical models is exhibited by the physical world, one may legitimately wonder why we can't be justified in attributing *independent* existence to the spacetime structure. There are at least four objections to the deflationary claim that I am attributing to Stein and trying to defend, the first three of which can be raised independently of GTR:

O_1 By playing the deflationary game, aren't we sweeping the philosophical problems under the carpet?

O_2 Stein's thesis depends on a controversial way of understanding the relationship between models and physical world. What does it mean, exactly, to claim that the world of physical events "*exhibits* a certain structure"?

[16] I am not presupposing here that Stein wanted to propose a *tertium quid* between substantivalism and relationism, let alone that he wanted to defend some form of what is now known as structural realism.

[17] This distinction has been pressed, among others, by Varzi (2001).

O_3 It is not at all meaningless or "supererogatory" to ask whether the space-time structure "really exists" in addition to its being exemplified.

O_4 Which entity does the exemplification of the structure, spacetime points or physical events/systems? If the former, Stein is wrong, if the latter Stein's SSR is pure relationism; in either case my reconstruction of his proposal does *not* amount to dissolving the substantivalism/relationism debate in GTR.[18] Let us discuss these four objections in turn.

As a response to O_1, consider the following analogy taken from the philosophy of time. Regarding *becoming as the successive occurring of events* accommodates both block-view theorists and the friends of becoming, depending on whether we insist on the fact that events *are* (static sounding) tenselessly located in spacetime, or on the fact that they *occur* (dynamic sounding) at their spacetime location.[19] In effect, since the *being* of events is identical with their *occurring*, we realize a fusion of Parmenideian and Heracliteian metaphysics.

Analogously, Stein's version of structural spacetime realism sounds *realist* about spacetime (and it is realist), because it claims that the physical world does indeed have a certain spatiotemporal structure (so in *this* restricted sense, spacetime exists), but it also sounds antirealist to those who keep asking the supererogatory question whether, in addition to characterizing the world in the specified manner, the "structure really exists". This solution to the substantivalist/relationist debate does not look like sweeping difficult questions under the carpet, but simply invites philosophers of space and time to deal with different problems.

Going to the second objection O_2, rather than implicitly defending the semantic view of theories, Stein explicitly advocates a "platonic", model-theoretic understanding of the relationship between mathematical models and physical world:

> «what I believe the history of science has shown is that on a certain very deep question, Aristotle was entirely wrong and Plato—at least on one reading, the one I prefer—remarkably right: namely, our science comes closest to comprehending the real, not in its account of "substances" and their kinds, but in its account of the "Forms" which phenomena "imitate" (for "Forms" read "theoretical structures," for "imitate" read "are represented by» (Stein, 1989, p. 52).

Here Stein's bent toward some of the tenets of structural realism is clear. The forms or "theoretical structures" are the mathematical, abstract models, which refer to the physical world by representing the relationships among those parts of physical systems that are described by laws. To the extent that a given physical process, say free fall, can be subsumed under a well-confirmed physical law, say the principle of equivalence, then one can "represent" that process by a *geometric* notion, that of a geodesic of a curved connection, which is part and parcel of the geometric structure of spacetime (for this view, see also DiSalle, 1995, p. 335). This structural realist way of construing the relationship between physics and geometry seems to me plausible and clear, and taking the notion of "the physical world

[18] The attentive reader will recall that this question had been raised in the previous section.

[19] For such a deflationary claim, see Savitt (2001), Dieks (2006), and Dorato (2006a, 2006b).

(free fall) exhibiting a certain geometric structure" as a primitive cannot be *prima facie* attacked for its inconsistency or lack of clarity.

The third objection O_3 affirms that, besides the hypothesis of manifold substantivalism, there are at least *three* different senses in which one could *meaningfully* ask whether spatiotemporal structure "really exists", in addition to being exemplified by the physical world. I will now argue that they are all supererogatory or irrelevant.

(1) In a first sense, the '*really* exists' in "the structure really exists" of the first quotation by Stein[20] could be taken as synonymous with 'mind-independently exist'. However, if we grant that spatio-temporal relations are exemplified by *physical* systems, who would want to deny their mind-independence? And even if one wanted to press the Kantian point that phenomena can be linked by spatiotemporal relations only thanks to our transcendental, pure intuitions of space and time, this rendering of the "really exists" would open a wholly different problem, not relevant to the one we started with.

(2) In a second sense, the "really exists" may refer to a kind of platonic realism about the mathematical structure used to model the physical world. This is a meaningful, *abstract* sense of "really exists", but also not relevant to our problem of establishing the *concrete* existence of spacetime.

(3) In a third sense, the question of the independent existence of spatiotemporal structure might call into play the ontic status of the truth makers of the equations defining the mathematical structure and expressing the laws of nature. *Via* the concept of symmetry, the spatiotemporal structure of spacetime is closely related to laws of nature, which in part codify and express such structure: granting the structure an independent existence might involve accepting a realist, possibly "necessitarist" position about laws of nature in the sense of Tooley–Dretske–Armstrong (see Earman, 1986). It must be admitted that this interpretation of "really exists" would not be meaningless, and that laws of nature, as opposed to laws of science, may indeed be attributed a primitive existence (Maudlin, 2007). However, questions concerning the metaphysical status of laws or the existence of universals *vis à vis* nominalistic interpretations of laws of nature involve *all* laws of nature, and not just those characterizing spacetime physics. As such, they do not seem specific enough for our gaining a deeper understanding of the ontological role of spacetime.[21]

Objection O_4 takes us closer to the interpretive problems of GTR, and seems the most threatening for my main argument. Given that spacetime is exemplified structure, one is naturally brought to ask what kinds of entities are the relata of the relations, so as to actually doing the "exemplificative work". If such an exemplification is realized by points of the manifold, we must assume their existence, as in manifold substantivalism; on the other hand, if it is realized by physical events/systems, we have relationism. Clearly, without additional arguments coming from GTR, structural spacetime realism, even in Stein's version, does *not* dissolve the debate.

[20] The one occurring just after the beginning of Section 2.

[21] Furthermore, in view of the remarks that will be offered in the next section, how do we distinguish laws involving the spatiotemporal structure from the other laws?

This is true, but note that this objection is predicated upon a clear distinction between spacetime and physical fields, a distinction which, as we are about to see in the next section, is definitely overcome by GTR. We will now see how also this fourth objection fails, and structural spacetime realism in the version defended here is vindicated.

3. THE DUAL ROLE OF THE METRIC FIELD IN GTR

As much as we have a particle-wave duality in QM, we have a (different) spacetime/physical field "duality" in GTR, forced upon us by the well-known dual role that the metric field has in the theory. As a matter of fact, *the metric field plays both the traditional roles represented by "space and time" and those typical of a physical entity*.

While, on the one hand, the metric field carries the distinction between spatial and temporal directions, allows measures of spatiotemporal distances, and specifies the inertial motions (as *geometric* entities typically do), on the other it also carries energy and momentum, satisfies differential equations, and acts upon matter, as *physical* fields do. The former roles leads us to claim that the metric field g_{ab} should be spacetime; the latter roles push us in the opposite direction, namely are conducive to maintain that it is the bare manifold that should represent spacetime, since the metric field is also, and indisputably, a physical entity. In reality, the tensor field g_{ab} has both roles, and I take it that this is the main, essential message of GTR. Since the metric field is both spacetime and a real, concrete physical field, *we should conclude that GTR is either both substantivalist and relationist, or neither substantivalist nor relationist*.

The question "which entity of the mathematical model should we regard as the representor of spacetime?" has, not surprisingly, generated *two* answers also in the literature, as it is illustrated also by the two available definitions of substantivalism provided in Section 1. Those who worried that g_{ab} is a physical field preferred to identify spacetime with whatever is denoted by the differentiable manifold, and thought that substantivalists are committed to *manifold substantivalism* (Earman and Norton, 1987; Earman, 1989; Belot and Earman, 2001; Saunders, 2003). Others, who correctly lamented that the manifold of events is deprived of any metric property, identified spacetime with the metric field plus the manifold (Maudlin, 1989; Stachel, 1993; Hoefer, 1996; Lusanna and Pauri 2006, 2007).

The fact that the candidate for representing "spacetime" has been oscillating between the manifold and the metric field is a first but important piece of evidence that in GR the debate lacks a clear formulation. This ambiguity, however, does not mean that our preference for regarding the metric rather than the manifold as representing spacetime is unmotivated. Even though I cannot rehearse the arguments in favor of this choice here, I will touch on three essential points, because they provide additional motivations to drop the substantivalism/relationism debate.[22]

[22] For additional arguments, I refer to the literature mentioned above. The invitation to drop the debate presupposes the context of our best, empirically confirmed spacetime theory so far, GTR.

The first is that we cannot even *talk* about "spacetime" without the resources provided by the metric, because in order to have spacetime, we need at least to be able to distinguish spatial from temporal intervals. Dimensionality alone, provided by the topological structure of the manifold, does not suffice.

In order to introduce the second argument, recall that it has been argued that if the metric field, rather than the manifold, becomes the "container", i.e., spacetime, then in those unified field theories *à la* Einstein, in which *any* kind of matter is represented by a generalized metric field, substantivalism would be trivialized. In such theories, in fact, there would be "nothing contained in spacetime", and substantivalism would amount to claiming the independent existence of the entire universe (Earman and Norton, 1987, p. 519). However, such an undesirable consequence can also be eliminated by dropping the substantivalism/relationism dichotomy altogether, at least to the extent that it implies a container/contained distinction. Why should we leave room for the meaningfulness of the latter distinction if the main point of GTR is to make spacetime a dynamic entity, capable of acting and reacting with the other matter fields? The dynamical character of spacetime, nevertheless, could seem to lend credibility to metric field substantivalism, and therefore to a form of *spacetime substantivalism* (Hoefer, 1996). If spacetime is the metric field and it is dynamical, why isn't it a substance?

The fact is that precisely because in GTR spacetime is *also* a physical entity, its role in the theory can always be redescribed by claiming that it is the manifestation of the gravitational field (its structural quality), rather than the other way around (the gravitational field being a manifestation of spacetime).[23] And the choice between these two ways of expressing the relationship between spacetime and gravitational field seems to be underdetermined by the facts, and suggests that the dispute between substantivalism and relationism in GTR is a matter of words, or possibly of a conventional choice about two ways of explaining phenomena that are empirically equivalent. If I claim that the gravitational field is a manifestation of spacetime, I start from the latter to "construe" the former, and I do the opposite in the reverse case, but both approaches look viable.

The third argument concerns the fact that all physical fields are assignment of properties to spacetime regions (Earman, 1989, pp. 158–159); so we should at least quantify over the points and regions of the differentiable manifold on which matter fields live. The reply is two-pronged; for non-gravitational matter, it is not clear why the points over which to quantify could not be those of the metric field, rather than the points of the manifold. Matter fields live on the metric field: as Rovelli once put it, "they live on top of each other". On the other hand, the question "where the points of the metric field are", if spacetime is the metric field or its structural quality, is clearly meaningless, as it would be equivalent to ask where is the universe, once we agree that universe (matter fields and gravitational fields) and spacetime are one and the same entity. In a word, also the Field's argument cannot go off the ground.

[23] In his abstract for the conference to be held in Montreal, Lehmkuhl (2006) has referred to these two alternatives as the fieldization of geometry and the geometrization of the field. He opts for a position that is very close to the one presented here.

Aware of these difficulties, Belot and Earman, who are convinced that the dispute between substantivalism and relationism still makes sense, put forward this account, which is equivalent to endorsing a metaphysics which is very close to heacceitism:

> «It is now somewhat more difficult to specify the nature of the disagreement between the two parties. It is no longer possible to cash out the disagreement in terms of the nature of absolute motion (absolute acceleration will be defined in terms of the four-dimensional geometrical structure that substantivalists and relationist *agree* about). We can however, still look at *possibilia* for a way of putting the issue. Some substantivalist, at least, will affirm, while all relationists will deny, that there are distinct possible world in which the same geometries are instantiated, but which are nonetheless distinct in virtue of the fact that different roles are played by different spacetime points (in this world, the maximum curvature occurs at this point, while it occurs at that point in the other world). We will call substantivalists who go along with these sorts of counterfactuals *straightforward substantivalists*. Not all substantivalists are straightforward: recent years have seen a proliferation of *sophisticated* substantivalist who ape relationists' denial of the relevant counterfactuals (Belot and Earman, 2001, p. 228).

If we regard as different two worlds that contain exactly the same individuals and properties, but vary only about which individual instantiate which properties, then we accept *haecceitism* (Lewis, 1986, p. 221). Imagine having two canvases (spacetimes), and to remove the content of the first picture from the first and paste it onto the second, in such a way as to shift it just by three inches to the left. The content of the two pictures is identical, only the second is moved to the left, and so different individuals (points) in the second canvas play different roles. Notice that in our example the frame allows for an independent identification of the points of the canvas, since the points in which, say, the flower is painted, have a different distance from the left, lowest corner.

In the example given by Belot and Earman, however, such an identification is impossible in principle, and not by chance they refer to the points by using an ostensive criterion (*this* point, or *that* point), and therefore presuppose an implicit reference frame, our bodies. The idea of a primitive thisness (heacceity) seems to stem from an identity criterion that is independent from anything pertaining to the causal role played by the individual or its properties. According to heacceitism, an individual is not the bundle of its properties, but, like a peg which can hold different clothes, has something substantial "under them", so that in an heacceitistic world *I* could have all *your* properties and keep my identity and *viceversa*.

This formulation of substantivalism is definitely supererogatory in Stein's sense. No possible *a posteriori* argument could ever be produced in favour of the kind of heacceitism that is required by the definition, since no empirical criterion whatsoever could in fact distinguish two physically possible worlds simply in virtue of the role played by the different points in the two models. And this result would be independent of the particular spacetime structure exemplified by the world of events, and would therefore be insensitive to the various types of

spacetime theories: the supererogatory nature of Belot–Earman approach to substantivalism is given by the fact that no possible a posteriori argument could ever be produced in favour of substantivalism/heacceitism.

Note, however, that this remark does not entail that in the context of GTR we should all become relationists. The metric field is spatiotemporal and physical at the same time, so that there is no clear sense in which we can distinguish physical entities from purely spatiotemporal relations, as relationism requires. The fact that also in the GTR case spacetime *is* exemplified structure does not entail that the metric field does not carry energy and momentum.

4. CONCLUSION

The metric field is spacetime, and it is a real entity, but the additional, metaphysical question whether it is a substance-like or relation-like is much less important than establishing its existence as exemplified structure, in the sense specified by structural spacetime realism. But structural spacetime realism turns into relationism only if we presuppose that the distinction between substantivalism and relationism has some utility in the philosophy of space and time.[24] However, as Newton had already understood, the categories of ordinary language (subject-predicate) as they have been re-elaborated by scholastic philosophy (substance-accident) seem quite inappropriate to understand the ontology of spacetime, or of any physical theory formulated in mathematical terms:

> «About extension, then, it is probably expected that it is being defined either as substance or accidents or nothing at all. But by no means nothing, surely, therefore it has some mode of existence proper to itself, by which it fits neither to substance nor to accident.» (Newton, 1685, p. 136)

If Newton, the alleged champion of substantivalism, argues that the notion of substance is "unintelligible" (see also DiSalle, 2002, p. 46), why using it after the invention of a theory (GTR) in which the distinction between container (spacetime) and contained (field) has evaporated?

ACKNOWLEDGEMENT

I am highly indebted to Michael Esfeld for his critical comments on a previous version of this chapter.

REFERENCES

Barrett, J.A., 1999. The Quantum Mechanics of Minds and Worlds. Oxford University Press, Oxford.

Belot, G., Earman, J., 2001. Pre-Socratic quantum gravity. In: Callender, C., Huggett, N. (Eds.), Philosophy Meets Physics at the Planck Scale. Cambridge University Press, pp. 213–255.

[24] For an historical reconstruction of spacetime theories that intentionally leaves on a side the question of substantivalism *vs.* relationism, see DiSalle (2006).

Cao, T.Y., 2003. Structural realism and the interpretation of quantum field theory. Synthese 1, 3–24.
Descartes, R., 1644. Principia Philosophiae. Transl. by B. Reynolds, Principles of Philosophy, 1988. E. Mellen Press, Lewiston, NY.
Dieks, D., 2006. Becoming, relativity and locality. In: Dieks, D. (Ed.), The Ontology of Spacetime. Elsevier, Amsterdam, pp. 157–176.
DiSalle, R., 1995. Spacetime theory as physical geometry. Erkenntnis 42, 317–337.
DiSalle, R., 2002. Newton's philosophical analysis of space and time. In: Cohen, I.B., Smith, G.E. (Eds.), The Cambridge Companion to Newton. Cambridge University Press, Cambridge, pp. 33–56.
DiSalle, R., 2006. Understanding Space-time. Cambridge University Press, Cambridge.
Dorato, M., 2000. Substantivalism, relationism, and structural spacetime realism. Foundations of Physics 30, 1605–1628.
Dorato, M., 2006a. Absolute becoming, relational becoming and the arrow of time: Some non-conventional remarks on the relationship between physics and metaphysics. Studies in History and Philosophy of Modern Physics 37 (3), 559–576.
Dorato, M., 2006b. Properties and dispositions: some metaphysical remarks on quantum ontology. In: Bassi, A., Dürr, D., Weber, T., Zanghì, N. (Eds.), Quantum Mechanics, Conference Proceedings. American Institute of Physics, Melville, New York, pp. 139–157.
Dorato, M., Pauri, M., 2006. Holism and structuralism in classical and quantum general relativity. In: Rickles, D., French, S., Saatsi, J. (Eds.), The Structural Foundations of Quantum Gravity. Oxford University Press, Oxford, pp. 121–151.
Earman, J., 1986. A Primer on Determinism. Reidel, Dordrecht.
Earman, J., Norton, J., 1987. What price spacetime substantivalism. British Journal for the Philosophy of Science 38, 515–525.
Earman, J., 1989. World enough and Space-Time: Absolute versus Relational Theories of Space and Time. The MIT Press, Cambridge, MA.
Einstein, A., 1961. Relativity and the problem of space. In: Relativity: The Special and the General Theory. Crown Publishers, Inc., New York.
Esfeld, M., 2004. Quantum entanglement and a metaphysics of relations. Studies in History and Philosophy of Modern Science 35B, 601–617.
Esfeld, M., Lam, V., 2006. Moderate structural realism about spacetime. Synthese, forthcoming. http://philsci-archive.pitt.edu/archive/00002778/.
Esfeld, M., 2007. Structures and powers. In: Alisa, O., Bokulich, P. (Eds.), Scientific Structuralism. Springer, Dordrecht, forthcoming.
French, S., Ladyman, J., 2003. Remodelling structural realism: quantum physics and the metaphysics of structure. Synthese 1, 31–56.
Hoefer, C., 1996. The metaphysics of space-time substantivalism. Journal of Philosophy 93, 5–27.
Janiak, A., 2004. Newton's Philosophical Writings. Cambridge University Press, Cambridge.
Ladyman, J., 1998. What is structural realism? Studies of History and Philosophy of Science 29 (3), 409–424.
Laudan, L., 1981. A confutation of convergent realism. Philosophy of Science 48, 19–49.
Langton, R., Lewis, D., 1998. Defining 'intrinsic'. Philosophy and Phenomenological Research 58, 333–345.
Lehmkuhl, D., 2006. Is spacetime a field? http://www.spacetimesociety.org/conferences/2006/docs/Lehmkuhl.pdf.
Lewis, D.K., 1986. On the Plurality of Worlds. Blackwell, Oxford.
Lusanna, L., Pauri, M., 2006. Explaining Leibniz equivalence as difference of non-inertial appearances: Dis-solution of the Hole Argument and physical individuation of point-events. Studies in History and Philosophy of Modern Physics, 692–725, arXiv: gr-qc/0503069.
Lusanna, L., Pauri, M., 2007. Dynamical emergence of instantaneous 3-spaces in a class of models of general relativity. In: Petkov, V. (Ed.), Relativity and Dimensionality of the World. Springer-Verlag, forthcoming. gr-qc/0611045. Pitt-Archive, IDcode:3032, 2006.
Maudlin, T., 1989. The essence of spacetime. In: Fine, A., Leplin, J. (Eds.), In: PSA 1988, vol. 2, pp. 82–91.
Maudlin, T., 2007. Metaphysics within Physics. Oxford University Press, Oxford.
Newton, I., 1685. De Gravitatione et equipondio fluidorum. In: Hall, A.R., Hall, M.B. (Eds.), Unpublished Scientific papers of Isaac Newton. Cambridge University Press, Cambridge, pp. 89–156.

Niiniluoto, I., 1999. Critical Scientific Realism. Clarendon Press, Oxford.
Poincaré, H., 1905. Science and Hypothesis. Walter Scott Publishing, London.
Pooley, O., 2006. Points, particles, and structural realism. In: Rickles, D., French, S., Saatsi, J. (Eds.), The Structural Foundations of Quantum Gravity. Oxford University Press, Oxford, pp. 83–120.
Psillos, S., 1999. Scientific Realism. How Science Tracks Truth. Routledge, London.
Rickles, D., French, S., 2006. Quantum gravity meets structuralism. In: Rickles, D., French, S., Saatsi, J. (Eds.), The Structural Foundations of Quantum Gravity. Oxford University Press, Oxford, pp. 1–39.
Rovelli, C., 1997. Half way through the woods. In: Earman, J., Norton, J. (Eds.), The Cosmos of Science. University of Pittsburgh Press, Universitäts Verlag Konstanz, Pittsburgh, Konstanz, pp. 180–223.
Rynasiewicz, R., 1996. Absolute versus relational space-time: An outmoded debate? Journal of Philosophy 43, 279–306.
Savitt, S., 2001. A limited defense of passage. American Philosophical Quarterly 38, 261–270.
Saunders, S., 2003. Indiscernibles, general covariance, and other symmetries. In: Ashtekar, A., Howard, D., Renn, J., Sarkar, S., Shimony, A. (Eds.), Revisiting the Foundations of Relativistic Physics: Festschrift in Honour of John Stachel. Kluwer, Dordrecht.
Slowik, E., 2006. Spacetime, ontology, and structural realism. http://philsci-archive.pitt.edu/archive/00002872/01/STSR2.doc.
Smith, P., 1998. Explaining Chaos. Cambridge University Press, Cambridge.
Spinoza, B., 2000. Ethics, edited and translated by G.H.R. Parkinson. Oxford University Press, Oxford, New York.
Stachel, J., 1993. The meaning of general covariance. In: Earman, J., et al. (Eds.), Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum. University of Pittsburgh Press, Universitätsverlag Konstanz, Pittsburgh, Konstanz, pp. 129–160.
Stachel, J., 2002. The relation between things versus the things between the relation: The deep meaning of the hole argument. In: Malament, D. (Ed.), Reading Natural Philosophy. Essays in the History and Philosophy of Science and Mathematics. Chicago University Press, Chicago, pp. 231–266.
Stein, H., 1967. Newtonian spacetime. Texas Quarterly 10, 174–200.
Stein, H., 1989. Yes, but… Some skeptical remarks on realism and antirealism. Dialectica 43, 47–65.
Suárez, M., 2004. Quantum selections, propensities, and the problem of measurement. British Journal for the Philosophy of Science 55, 219–255.
Van Fraassen, B., 1980. The Scientific Image. Oxford University Press, Oxford.
Varzi, A., 2001. Parole, oggetti eventi. Carocci, Roma.
Worrall, J., 1989. Structural realism: The best of both worlds? Dialectica 43, 99–124.

CHAPTER 3

Identity, Spacetime, and Cosmology

Jan Faye*

Abstract

Modern cosmology treats space and time, or rather space-time, as concrete particulars. The General Theory of Relativity combines the distribution of matter and energy with the curvature of space-time. Here space-time appears as a concrete entity which affects matter and energy and is affected by the things in it. I question the idea that space-time is a concrete existing entity, which both substantivalism and reductive relationism maintain. Instead I propose an alternative view, which may be called non-reductive relationism, by arguing that space and time are abstract entities based on extension and changes.

Theories about the nature of space and time come traditionally in two versions. Some regard space and time to be substantival in the sense that they consider space-time points fundamental entities in their own right; others take space and time to be relational by somehow constructing points and moments out of objects and events. In spite of their fundamental disagreements, substantivalists and relationists share a common view: they regard space and time as concrete particulars. Hence Quine's famous dictum "no entity without identity" should apply to space and time. Supposing the existence of a concrete particular, we must be able to point to some determinate identity conditions of space and time which would allow us to regard them as concrete particulars. In fact, most philosophers just take for granted that space and time are concrete entities; they tacitly presume that appropriate identity conditions exist and that it is rather unproblematic to specify what these are.

In his seminal work on the debate about absolute and relational theories of space and time, Earman (1989) points to the serious difficulties concerning identity and individuation any theory of space-point must confront. After having discussed various metaphysical accounts of predication, he makes the following remarks:

* Department of Media, Cognition, and Communication University of Copenhagen, Denmark

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00003-X
© Elsevier BV
All rights reserved

> One could try to escape these difficulties by saying of space-time points what has been said of the natural numbers, namely, that they are abstract rather than concrete objects in that they are to be identified with an order type. But this escape route robs space-time points of much of their substantiality and thus renders obscure the meaning of physical determinism understood, as the substantivalist would have it, as a doctrine about the uniqueness of the unfolding of events at space-time locations. (p. 199)

Earman does not develop this suggestion because, as he observes, it deviates too much from the substantivalist core assumptions. I shall, however, attempt to lay out a view according to which space-time is taken to be an abstract entity. But first I shall review some of the difficulties which Earman mentions in the light of recent discussions on identity and individuality.

Until recently I shared the concreteness view of space and time, or spacetime. But I couldn't find any plausible identity conditions for space and time, or spacetime, to be concrete particulars. I have since begun to think that spacetime points should be categorized as abstract particulars.[1] I don't know whether this puts me in bad company, but I think Leibniz meant something similar in his correspondence with Clarke when he pointed out that space and time are not fully real but are 'ideals'.[2] Space, I submit, refers to the set of all bodies, and time designates the set of all changes with a beginning and an end. I believe that this position has some very important explanatory advantages and that it may even open up possibilities for a satisfactory solution to the debate between the relationist and the substantivalist. I shall present some arguments to the effect that space and time, or spacetime, should be considered abstract particulars. By this I mean that locations and moments are existent things—abstract man-made artifacts whose role is to help us represent the world and thereby identify and individuate concrete objects.

My suggestion is that space-time is an abstract object whose structure supervenes on actual things and events. For the sake of terminological clarification, I take an abstract object to be an entity which is existentially dependent on concrete particulars that instantiate it and whose identity does not fulfil the normal determinate identity condition of concrete objects. Similarly, I take particular properties and relations to be tropes that are instances of universals. While I recognize that other philosophers use these terms differently, lack of space prohibits me from discussing those uses here.

THE EXISTENCE OF SPACE

Whether we think of space as being absolute or relational, space is considered to be physically real. Either view takes for granted that spatial locations exist. The absolutist, in being a substantivalist, believes that space points exist over and

[1] In an earlier paper (Faye, 2006b), I argued that time is an abstract entity but kept a door open for the concreteness of space. Also I counted Leibniz as a proponent of space and time as concretes because I took him for being a reductionist by heart. Now, having reconsidered, I must admit that this remark may be too hasty.

[2] Indeed, 'ideal' have several meanings. By using 'ideal' in contrast to 'real', Leibniz seems to think of space and time as something whose existence (partly) depends on the mind.

above what is located in them, and that these points have intrinsic relations to one another. The relationist, in contrast, argues that spatial locations are nothing by themselves, since they are reducible to things that are said to occupy them, and which can be directly related by physical processes among these things.

Historically these characterizations may not be true of the two arch contestants of substantivalism and relationism respectively. Newton denied that space and time are real substances; nor did he think that they are accidents. He seems to have taken over Pierre Gassendi's view that space and time are of a third kind, claiming that space and time are preconditions of substance. Before Newton, Gassendi argued that time flows uniformly regardless of any motion, and that space is uniformly extended irrespectively of the bodies it may contain.[3] Newton associated space and time with modes of existence because of his assumption of God as the necessary Being who is substantially omnipresent and eternal. Nonetheless, he claimed: "Although space may be empty of body, nevertheless it is not itself a void; and *something* is there because spaces are there, though nothing more than that" (Hall and Hall, 1962: 138). He also emphasized that space is distinct from body and that bodies fill the space. So Newton seems to be as close to being a substantivalist as one can be, especially if one brackets his belief in God and consider space to be an immaterial substance which can exist empty of any material substances. Similarly, Leibniz was less of a hard-core relationist than was Descartes. In his correspondence with Clarke, he explicitly said that space and time are 'ideals', having no full reality. Space "being neither a substance, nor an accident, it must be a mere ideal thing, the consideration of which is nevertheless useful" (Alexander, 1956: 71). This is interesting because it indicates that Leibniz saw space and time as abstractions rather than an aggregate of spatial and temporal relations between concrete objects.

If locations or space points are concrete entities, it must be possible to specify their identity conditions in a way showing that they are concrete entities. Space points are in Space, and being in Space is a criterion of being a concrete entity. However, space points cannot exist independently of Space itself. Particular locations are intrinsically featureless; they lack any internal features for differentiation among themselves. Being parts of Space they have, by necessity, the same nature of identity as Space itself in terms of being concrete or abstract. Space is not just the mereological sum of its parts even though spaces may seem to be absolutely the same all the way down because Space, taken to be a substance, contains an absolute metric that cannot emerge from a collection of individual points. Rather the individuality of the points comes from the structure of Space itself. Bearing witness to this claim, Newton said:

> The parts of duration and space are only understood to be the same as they really are because of their mutual order and position, nor do they have any hint of individuality apart from that order and position which consequently cannot be altered. (Hall and Hall, 1962: 136)

A location depends for its existence upon Space and consequently its identity depends on the identity of Space. Thus, if locations are concrete particulars, then

[3] See Gassendi (1972), p. 383 ff.

Space itself must be a concrete particular. But Space itself cannot be in space, because that would make Space a part of space; thus its identity would depend on this further space. Therefore space points cannot be concrete entities.

Nor does the causal criterion of concreteness apply to Space. Although it has been held that Newton considered absolute space to be a cause of the inertial forces, there is no textual evidence for such an interpretation, and it seems more accurate to say that Newton believed that absolute space merely acts as a frame of reference and that acceleration by itself gives rise to the inertial forces. The relationist, however, hopes to account for the distinction between relative motion and 'real' accelerated motion not in terms of absolute space, or any other object to which the motion is relative, but in terms of causes of the motion.

A third conception of abstract entities takes them to be incapable of existing independent of other things.[4] We may define a substance as a concrete particular whose existence does not depend for its existence on any other particular. It then follows, by contrast, that a particular whose existence is dependent on other particulars is an abstract entity in the sense under discussion. Indeed, it may be possible in thought to separate two particulars where one existentially depends on the other. An illustration of such a separation would be whenever we think of a particular colour (a trope) as being divided from the object which it is a colour of. Nevertheless, this view seems to exclude events from being concrete particulars since they exist inseparately from those things they involve. The emission of light cannot exist independently of the source which produces it. But events are concrete particulars to the extent that they exist in space and time, they also partake in causal explanations, and sometimes we even identify a concrete object in virtue of a certain event. A sudden flare on the sky, a supernova, may be used to identify the star that once has exploded. So an entity can be an abstract one in the sense of being existentially dependent upon other entities but still be pointed to as a concrete object in terms of having a location in space and time.

Also we find that locations of things can be defined in terms of functional expressions such as 'The location of Montreal' is the same as 'The location of the largest city of Canada', where the identity conditions of locations is dependent on things occupying them and the spatial relations. At first glance it seems possible to identify locations quite independently of the physical things.

- $(x)(y)((x = y)$ if and only if space(x) & space(y) & x coincides with $y)$.

But we have just learned that the individuality of locations depends on the order of Space itself; hence if Space is not a concrete object, neither can locations be. Furthermore, we should notice that the relation 'coincides with' is reflexive, symmetric, and transitive as required by the abstraction principle. Locations can therefore be pointed out in relation to concrete particulars and their mutual spatial relations.

The proper identity condition for locations is then expressed by a proposition which grounds the abstract sortal term 'location' in the coincident relation between things or other concrete particulars:

- $(x)(y)((\mathrm{loc}\, x = \mathrm{loc}\, y)$ if and only if thing(x) & thing(y) & x coincides with $y)$.

[4] See, for instance, Lowe (1998), Ch. 10.

Neat as the statement seems, it is nonetheless obvious that it negates the existence of empty space. Avoiding any animosity of the void (between separated things) we must allow a modal formulation like the following:

- $(x)(y)((\mathrm{loc}\, x = \mathrm{loc}\, y)$ if and only if thing(x) & thing(y) & 1) x coincides with y, or 2) in case y and y did not exist, then if they had existed, x would have coincided with y whenever y would have coincided with x, and vice versa).

This illustrates that locations are actually distinct from physically things but still logically incapable of existing separately from physical things as such.

THE EXISTENCE OF TIME

Aristotle said that time is not change but the measure of change, or rather "that in respect of which change is numerable" (Aristoteles, 1955, *Physics*, 219^{b}2). This suggestion was perhaps not such a bad choice. Motion is something we can perceive. Together with extension in space, change and motion are what we can immediately see by the naked eye, whereas space and time is what we apparently only are able to see indirectly with the help of the celestial motion of objects such as the sun or mechanical clocks. But there is more to Aristotle's suggestion than epistemological priority. In addition, his remark implies that our understanding of motion is prior to that of time. Also time *is* nothing but a measure of motion. Given this interpretation, motion is not only semantically prior but likewise ontologically prior to time. The existence of motion and change precedes the existence of time. Aristotle's ontology of time thus comes close to our everyday experience of temporality.

This also explains why Aristotle seems to defy the existence of time instants. He says in connection with Zeno's paradoxes: Zeno's conclusion "follows from the assumption that time is composed of moments: if this assumption is not granted, the conclusion will not follow" (Aristoteles, 1955, *Physics* 239^{b}30-3). What Aristotle probably had in mind was something like this: Since time is continuous, then each period of time must contain an infinite number of instants. But, according to him, nothing is *actually* infinite, but only *potentially* infinite. Numbers are in this way infinite in so far as there is no limit built into the process of counting. Likewise we can divide a length or a period of time in as many points or instants as we want, there is no limit to such divisions, but the divisions do not exist independently of the one who makes them. Hence, the potentially infinite divisibility does not imply the existence of actually infinite divisibility, and therefore spatial points and temporal instants do not exist independently of us. Although Aristotle does not explicitly say so, his view is not so far from saying that points and moments are not concrete entities but abstracted ones being the product of the converging limit of our cognitive ability to divide things up in smaller and smaller regions and intervals.

Following up on Aristotle, we may say that space and time cannot exist as a measure of motion unless things in motion exist prior to the numbering. Time exists only if change and motion exist. It is impossible for time to exist in case there is no change or motion. Thus, we see here an exemplification of the conception

of abstractness according to which existential dependence marks what it means to be an abstract entity; time ontologically depends on things in motion or things which undergo change. Moments are abstracted from varying things but do not exist independently of the concrete particulars from which they are abstracted. In contrast, substantivalism—as we find it in Newton's notion of absolute space and time or in a realist interpretation of Einstein's theory of space-time—takes moments to be ontologically prior to those physical events that may occupy them; time, or space-time, exists as an independent entity, whereas reductionism regards moments to be identical to physical events or their existence to be somehow parasitic on things and processes. Both views consider time to be a concrete particular. The first view captures time as a substance, the second view as a non-substance. This means that it must be possible to specify some identity criteria which show that time is a concrete particular. But what are they?

Moments or temporal instants seem to be concrete particulars existing in time because they stand in temporal relations to other times, and we seem to have no problems of specifying identity criteria for moments. We say:

- $(t)(t^*)((t = t^*)$ if and only if time(t) & time(t^*) & t is simultaneous with $t^*)$.

But time instants cannot exist independently of Time itself; they are parts of Time, and as parts of Time they must have the same nature of identity as Time itself in form of being concrete or being abstract. A temporal instant depends for its existence upon Time, which implies that the identity of a temporal instant depends on the identity of Time. Therefore we must expect Time to be a concrete particular because if moments are concrete particulars, then Time itself must be a concrete particular.

Assume that Time is a substance. Time should then, like any other physical substance, exist in space and time. But Time does not exist in time, whereas Space may be said to exist in time; thus space and time cannot determine the identity of Time. Hence Time cannot be an individual substance (Faye, 2006a, 2006b). Assume, in contrast, that Time is a non-substance because all talk about moments and temporal relations can be reduced to talks about events and causal relations. This requires that we can set up identity criteria of events which avoid any reference to space and time. Here Davidson's attempt to specify determinate identity criteria of events in terms of causation comes to mind as the only serious suggestion, claiming that:

- $(x)(y)((x = y)$ if and only if event(x) & event(y) & x and y cause and are caused by the same events).

Unfortunately the criterion has rightly and often been charged as being circular (Faye, 1989: 153–160). Thus, the conclusion seems to be inescapable. Time cannot be a concrete entity.

In contrast, I propose that the concept of time is an abstraction in the sense that we have a constructed temporal language to be able to talk about collections or sets of concrete changes. Time denotes a tenselessly ordered set of all events in the world. This suggestion is supported by the above conceptions of abstractness. Time does not exist in space and time. Again, time does not have any causal influ-

ence on concrete substances because, if it had had such an influence, then each and every particular event would be causally overdetermined by causally prior events and by the definite moment at which the event takes place since both the causally prior events and the moment in question would be causally sufficient for it. Time instants are therefore causally superfluous. Moreover, if we think of two events, which are causally connected so that the cause is not only causally sufficient, but also causally necessary for the event, there is no room for causally active moments.

Facing the third criterion of abstractness, we see that Time, like events, is logically incapable of existing separate from particular substances. Events, however, in contrast to Time, do exist in space and time, and thus we shall leave aside that they may be abstracts in some other sense. Time cannot exist without changing things; nevertheless we can, of course, separate time from substances in thought.

Finally, the concept of a temporal instant fulfils the principle of abstraction. We can assign a time instant to an event in terms of a functional expression and thus express the identity of moments in terms of identity of events. We say, for instance, the time of the Big Bang, the time of the supernova, and the time of the solar eclipse. These functional expressions meet the abstraction principles.

- $(x)(y)((\text{inst}\,x = \text{inst}\,y)$ if and only if event(x) & event(y) & x coexists with y in a frame S).

It says that the moment of x and the moment of y are identical if and only if x and y are events, and x and y coexist. The relation 'coexistence' is indeed reflexive, symmetric, and transitive in any given inertial frame, and it grounds the abstract sortal term 'instant' such that the understanding of instants or moments presupposes an understanding of events and changes.

SPACE-TIME SUBSTANTIVALISM

Up to now we have mainly considered whether space and time are concrete or abstract entities in a metaphysical context. Let us go on to consider the matter in a physical context.

In modern physics space and time merge into a single dynamic entity called space-time. It is sometimes assumed that this entity, according to the field equations of the general theory of relativity, is causally efficacious in the sense that space-time causes the distribution of matter and energy in the universe which in return affects the curvature of space-time. This assumption of mutual influence requires, being true, that space-time is a concrete entity which is able to undergo changes that effect or are affected by changes in the matter and energy distribution. However, changes take place only in something which exists persistently through these changes. If one accepts this metaphysical principle, it leads to the conclusion that space-time should be treated as an object, or rather a substance, which forms the persistent ground for any change. Therefore the assumption that space-time can be causally influenced, or can causally influence, presupposes substantivalism of some sort. Space-time is a real substance undergoing changes but which exists independently of those processes occurring within space-time.

Indeed, space-time substantivalism constitutes a serious threat to the claim that space and time are abstracts because it considers space-time points as concrete particulars. The proponents point out that the general theory of relativity quantifies over space-time points, and as true followers of Quine they take this as a reason for believing in the existence of space-time points. We therefore need to take a closer look at this view.

In their joint paper, Earman and Norton (1987) define 'substantivalism' as the claim that space-time has an identity independent of the fields contained in it. They emphasize that the equations describing these fields "are simply not sufficiently strong to determine uniquely all the spatio-temporal properties to which the substantivalist is committed" (Earman and Norton, 1987: 516). This catches the standard view that a substance is something that is self-subsistent. We may define a substance as a particular whose identity does not depend on any other particular, and whose existence therefore does not depend on it.

Before we proceed an important distinction should be made between *manifold substantivalism* and *metric substantivalism*. The first type forms a sort of minimal view according to which space-time consists of a topological manifold of points, and the metric field is then attached as an externally defined field, whereas the second type includes the metric field as an intrinsic part of the container itself.

Earman and Norton identify space-time with space-time points. As they say: "Thus we look upon the bare manifold—the 'container' of these fields—as space-time" (1987: 518–9). The bare manifold consists of space-time points, whereas the fields form the metrical structure of space-time which is added to the manifold as a thing in it. Their motivation for separating the bare manifold as the space-time container and the metric fields as the contained is that the metric fields carry energy and momentum which can be converted into other forms of energy and heat.

Manifold substantivalism takes space-time points to be real, but it is entirely unclear what their identity conditions are. It has been noted before that according to Newton the parts of absolute space and time are intrinsically identical to one another and can only be differentiated by their mutual intrinsic order.[5] But this move is foreclosed to the manifold substantivalist. The identity conditions for space-time points cannot involve the metrical structure because of the way manifold substantivalism has been defined. It is assumed that space-time is nothing over and above space-time points in the sense that the identity of the space-time manifold is dependent on the identity of space-time points. The consequence is that space-time is a real concrete particular if and only if space-time points have an independent identity.

Nevertheless, it appears reasonable to say that space-time is not a composite substance because the whole does not distinguish itself from the parts. Space-time is indefinitely divisible into other particulars of the same kind, but how can we distinguish between these parts in such a way that the distinction represents a real difference? Establishing determinate identity conditions, which make space-time a concrete entity, is a serious problem for manifold substantivalism. The points of

[5] See also T. Maudlin (1989), p. 86.

the manifold are pure abstract individuals, bare mathematical particulars which do not have any structure or properties in virtue of itself. How can we make sure that these mathematical objects represent self-subsistent physical space-time points (or real events)?

The manifold substantivalist seems to have two possibilities for formulating determinate identity conditions of space-time points. He can either follow the mathematical road or take the physical one. After all he may regard geometrical points as names of physical points, or he can insist on some form of metric structure as belonging to space-time (because we are able to talk about a universe free of matter and energy).

Following the first path, the manifold substantivalist does not bump into the concrete structure of the world, but the road is nonetheless not passable. Because it is impossible to see how geometrical points can act as names for physical space-time points, unless we already possess independent physical criteria of individuating space-time points. A name refers to what it names; but it can only be assigned to the named, in case the name and the named have mutually independent identity conditions. The manifold substantivalist, however, is unable to point to what these are with respect to physical space-time points. Mathematical points are all we have, and they have no intrinsic features that individuate them from each other.

As we have seen, space-time points can only be defined relatively within a relational structure, and their *only* identity is given in virtue of their position in this structure. We may indeed assign coordinates to the manifold. But in a pure differential manifold each and every possible form of coordination is arbitrary and the manifold is invariant with respect to the choice of a particular coordinate system. Only by adding a structure is it possible to change the situation, but then we no longer are confronting a bare manifold.

Choosing the second road, the manifold substantivalist may seek the identity conditions of space-time points in the metric structure of the physical state of the universe (*versus* Earman and Norton). In this way he may attempt to uphold a view of space-time as a concrete entity. If space-time is taken to be represented by a manifold of geometrical points on which we define a metric field, then the set of physical space-time points is individuated by their metric properties as they are defined by our best space-time theories.

In the theory of general relativity the metric field is identified with the gravitational field and it therefore carries momentum and energy. Let me quote John Norton:

> "This energy and momentum is freely interchanged with other matter fields in space-times. It is the source of the huge quantities of energy released as radiation and heat in stellar collapse, for example. To carry energy and momentum is a natural distinguishing characteristic of matter contained within space-time. So the metric field of general relativity seems to defy easy characterization. We would like it to be exclusively part of space-time the container, or exclusively part of matter the contained. Yet is seems to be part of both." (Norton, 2004)

Indeed, if the energy and momentum of the gravitational field can be converted into radiation and heat, and vice versa, in connection with the formation of back holes, and this field also characterized the metric properties of space-time, how can space-time exist independently of what is going on in it? Because the identity of space-time points logically depends on their metrical structure, they are incapable of existing without this structure.

The manifold substantivalist may respond by pointing out that Einstein's field equations connect the intrinsic structure of space-time with the distribution of matter and energy such that the metric field, in the form of the gravitational mass-energy field, and the matter field stand in a causal relationship. Thus, if space-time had no momentum and energy, it would be impossible to see how they could interact with matter. Moreover, we can only have a causal relation in case the relata are logically distinct from one another; i.e., in case the relata have mutually independent identity conditions. Thus, if space-time and stars and galaxies were separate entities, then their mutually causal interaction would constitute the proof that they are concrete particulars. But the argument, as it stands, is not without problems.

I sympathize with Lawrence Sklar as he warns us about believing that the field equations should be interpreted as the non-gravitational mass-energy causing modifications of space-time since "the possible distribution of mass-energy throughout a spacetime depends upon the intrinsic geometry of that spacetime." (Sklar, 1974: 75) Apparently, what he wants to emphasize is that the matter field is spatially and temporally distributed, and thus it cannot gain the necessary ontological independence of the metric field which is required of it in order to have a separate existence as necessary for causal efficiency. Instead, Sklar maintains that the equation should be interpreted as a law of coexistence:

> The equation tells us that given *both* a certain intrinsic geometry for spacetime and a specification of the distribution of mass-energy throughout this spacetime, the joint description is the description of a general-relativistically possible world only if the two descriptions jointly obey the field equation. (Sklar, 1974: 75)

Such a law-like constraint on the two descriptions robs the substantivalist of the causal argument for space-time and the matter field being concrete, independent particulars.

Where does this take us? It seems that manifold substantivalism either is forced into an abstract mathematical entity (since space-time points become abstract particulars) or collapses into a form of relationism where space-time as such is claimed to be identical with the fields of gravitation-cum-matter. In the latter case the metric field is defined in terms of the gravitational fields whereas the space-time points are defined in terms of the mass-energy fields. So manifold substantivalism seems not to be a viable metaphysical possibility if one wants to sustain a claim that space-time is a concrete substance.

In the debate about manifold substantivalism, according to which space-time is represented by a manifold of points and a metric field is added to the manifold, one argument appears to be more prominent than any other: the hole argument. It

apparently shows that a substantivalist interpretation of space-time requires that we are willing to ascribe a surplus of properties to space-time which is impossible for observation or the laws of the relevant space-time theories to determine. The substantivalist must concede that matter fields, which after a transformation go through such a hole in the space-time manifold, are not determined by the metric fields and the matter fields outside the hole. Nevertheless the manifold substantivalist, who wants to save determinism, also holds that there has to be physical differences between the possible trajectories which a galaxy may take inside the hole. Earman and Norton take this to be a most unwelcome consequence of space-time substantivalism and are ready to give up manifold substantivalism as such (Earman, 1989: Ch. 9). Attempts to avoid such a conclusion by adding further structure to the manifold can, at least in some important cases, be met by alternative versions of the hole argument (Norton, 1988). If manifold substantivalism has to give away, Earman sees three ways to uphold substantivalism with respect to space-time points. One may adopt a structural role theory of identity of space-time point (which I shall return to below in the form of *sophisticated substantivalism*), one may argue that metrical properties are essential to space-time points (Maudlin 1989, 1990), or one may introduce counterpart theory to space-time models (Butterfield, 1989). But in conclusion he finds that "our initial survey of the possibilities was not encouraging" (Earman, 1989: 207–208).

The central claim of metric substantivalism, according to Maudlin, is that "Physical space-time regions cannot exist without, and maintain no identity apart from, the particular spatio-temporal relations which obtain between them" (Maudlin, 1990: 545). Thus, the identity conditions of space-time points are determined by the intrinsic order among them. A few pages later he states that space-time and metric are connected by necessity: "Since space-time has its spatio-temporal features essentially (cf. Newton above), the metric is essential to it and matter fields not" (Maudlin, 1990: 547). The proponent of the metric substantivalism, in contrast to manifold substantivalism, welcomes the idea that space-time carries energy in the form of its metrical structure because it makes space-time on a par with other substances.[6]

In the general theory of relativity the metric field is associated with the gravitational field because of the proportionality of the gravitational and inertial mass so that gravitation and accelerated coordinate systems can be considered physically equivalent. Einstein spoke about this association in various terms: The gravitational field is said to either *influence* (or determine) or *define* the metrical properties of space-time.[7] But holding that the gravitational field *defines* the metric structure of space time, it must be an essential feature of the universe and not just accidental that gravitational and inertial mass is proportional. This indicates, of course, that the proportionality is due to the fact that gravitational field is logically identical to the metric field. Another possibility is to think of them as conceptually distinct but

[6] For a discussion of this argument, see Hoefer (2000).

[7] In his introduction to the Leibniz–Clare Correspondence, Alexander (1956), p. liv states two quotations of Einstein without any references, one in which Einstein says that the gravitational field 'influences or even determines the metric laws of the space-time continuum', the other in which he maintains that the gravitational fields 'define the metrical properties of the space measured'. The first is from Einstein (1955, p. 62), whereas the second has not been possible to locate.

empirically identical. However, according to Kripke, if such an identity proposition is true, it is necessarily true.

Very few, I believe, would argue that inertia and gravitation are not conceptually distinct. But when the intrinsic geometry of space-time is *identified* with the structure of the gravitation field, it cannot be an empirical discovery similar to the one that Hesperus and Phosphorus are the same. To see this we should realize what it takes to be an empirical discovery. It means that observation brings together evidence that fulfils two different identifying descriptions. Ancient astronomers possessed different, empirically based, criteria of being Hesperus and of being Phosphorus. But when it comes to identifying the metric of space-time with the gravitational field there are within GRT no such empirically based independent criteria of being a definite metric structure apart from the gravitational field itself. We should also remember that the equivalence of the gravitational field and accelerated frames is merely local. This gives us problems with a global assignment of a unique metric structure founded on the gravitational field. Second, what the association of the gravitational field with the metric structure of spacetime itself does is that we physically narrow down which of the possible abstract space-time models can be the model of the actual world. So the association is not an empirical identification but a metaphysical assumption that allows us to ground space-time talks in physical reality.

Indeed, there is a sense in which inertia and gravitation are the same property that is only described in two different ways in different frames. The principle of equivalence ensures an explanation of the proportionality between the gravitational mass and the inertial mass because it tells us that a system in free fall is an inertial system (locally). Therefore, it is a widely spread understanding of GRT that the metric field (or together with some related geometrical objects like connection…) represents both the space-time geometry and the gravitational field. So when it is said that it has been *decided* by the physics community that it is meaningful to identify the gravitational field with the metric field such a decision must be based on some assumption which is not an empirical discovery.[8] Rather the decision is based on a metaphysical assumption of co-existence according to which it is physical impossible that the metric field can exist independently of the gravitational field.

This brings me to the second part of the argument. Maudlin considers the metric field as an essential part of space-time substantivalism. As we have just seen, the metric structure of space-time is connected by necessity to the gravitational field where the notion of necessity is to be understood in a metaphysical sense and not merely in a physical sense.[9] Thus space-time is an entity whose existence cannot be separated from the existence of the gravitation. Space-time points and the metric structures we assign to these points are geometrical abstracts.

Assuming this is correct, it is metaphysically impossible for space-time to exist separable from gravitation. I therefore think that the four-dimensional represen-

[8] A point made by an anonymous referee.

[9] When Maudlin (1990) argues that "The substantivalist can regard the field equation as contingent truths, so that it is metaphysically possible for a particularly curved space-time to exist even if all of the matter in it were annihilated" (p. 551), he is talking about something else. Even if all matter is annihilated there still exists a so-called source free gravitational field which constitutes the metric field (see Norton, 1985: 243–244).

tation of the world is an abstraction. Such an abstracted entity as space-time with a metric and a topology is rich in structure and it therefore helps us to grasp a changing reality.

SPACE-TIME RELATIONISM

The proponent of the concreteness of space-time is not limited to substantivalism. He could still argue that space-time is a real entity as it reduces to spatiotemporal relations among the galaxies in the universe. But how can space-time points be concrete individuals without being a substance? The argument goes that space-time points are concrete because they owe their identity to concrete objects which occupy space and time. Especially they owe their identity to continuants or rather physical events.

Relationism, however, does not fare any better than substantivalism. I shall not rehearse all the kinematical-dynamical arguments which have been put against it by Sklar, Friedman, Earman, and others. What is important for my purposes is that the relationist believes that space-time does not exist over and above the concrete fields; he sees it merely as 'a structural quality of the field,' and therefore claims that all talk about space-time points reduces to talks about a causal-equivalence class of events. By this founding manoeuvre the relationist regards space-time talks as concerned with concrete particulars as much as the substantivalist does. But the relationist's attempt to specify such an equivalent class of causally connected events suffers from the lack of a consistent criterion of identity which leaves out space-time points.

The claim is that space-time points exist whenever events that occupy them exist. Thus space-time points are concrete because they reduce to concrete events in them. Space and time are identical to the things and events which are supposedly 'in' space-time. Events are then really constitutive parts of space-time analogous to the way our arms and legs are not in our body, but parts of it, i.e. constitutive parts. I think, however, that this escape route is no way out.

I suggest that we can only have ontological reduction in case a certain identity relation exists between the entity, which we want to reduce, and the parts to which we want to reduce it. The parts of a whole must not be exchangeable without the whole losing its identity. Thus, if a particular entity continues to be the same even if parts of it are replaced by different entities because the identity of such an entity is not dependent on the identity of the parts, then this entity is not reducible to the sum of its parts. An example: a human body does not consist of the mereological sum of its parts because the various organs may be transplanted by donor organs or artificial parts without the body discontinuing being the same. In contrast, however, particulars like particular masses or quantities of stuff are numerically the same as the sum of their parts because they depend for their identity upon the identities of objects which are their own proper parts.

Although impossible to perform it seems correct to say that a planet, the sun or a galaxy could be replaced by another object of its kind without space-time changing its identity. Space-time would still have the same curvature everywhere

and at every time, it would have the same metrical structure due to the same field of gravitation, and it would still be a four-dimensional continuum. It seems at least that all individual objects can be substituted by other material objects whereas the intrinsic properties of space-time, which ground the identity condition of space-time, stay the same all the way through.

Indeed, there are less radical forms of relationism. One can argue: 1) that space-time points exist only in virtue of those continuants and events which occupy them even though they are ontologically distinct from them; or 2) that space-time points exist only as possible places for continuants and events to exist. The metaphysical basis of the first claim is that an entity can be ontologically distinct from another entity only if they have independent identity conditions (as father and son). By making the identity of space-time points distinct from the identity of their occupants but by claiming them to be existentially dependent on these occupants, we do only make a separation in thought because their acclaimed distinct identity conditions do not have empirical consequences. This view collapses, in my opinion, to a claim that space-time points are abstract entities. The second claim, however, presupposes in contrast that the possible places have some kind of independent existence of their occupants. This view therefore gives a way for a sort of substantivalism. Thus none of the other forms of relationism do better than the radical one and save space-time points as concrete entities.

SPACE-TIME AS AN ABSTRACT ENTITY

In my opinion, the traditional dichotomy between substantivalism and relationism is false: Either (a) space-time is an ontologically independent entity because it can exist independently of physical things or events, or (b) it is reducible to structural properties of things or events. But substantivalism and relationism are not contradictory terms. (a) implies that things or events are not *necessary* for space and time; whereas (b) implies that events or things are *sufficient* for space and time by presupposing that things and events are definable or identifiable without any reference to space and time. (b) expresses only *reductive relationism*, and one can easily deny (a) without being committed to (b). Things and events can be *necessary* conditions for space and time even though space and time cannot non-circularly be reduced to things and events. I want to argue that space and time can be understood as abstracted from certain structural property of the physical world, and as such space-time is an abstract representation of these things and events. Geometry and pure theories of space and time in general are logical or mathematical abstractions from a physical implementation, but it is a serious mistake, I think, to hypostatize these abstractions. This view I call *non-reductive relationism*.

Non-reductive relationism takes the metric tensor g to represent a gravitational field rather than the space-time structure itself.[10] Field theories seem to change the

[10] Carlo Rovelli (1997: 193–94) argues that Einstein's identification between gravitational field and geometry can be understood in opposite ways: 1) "the gravitational field is nothing but a local distortion of spacetime geometry"; or 2) "*spacetime geometry is nothing but a manifestation of a particular physical field*, the gravitational field." He himself defends the second option which I take to be an example of reductive relationism. The metric field is the manifestation of the

long-established debate between Newton and Leibniz. The non-reductive relationist does not have to fight the notion of empty space. There is no space where there are no fields, i.e. something physical. The attempt to maintain the classical perspective by defining the physical matter in terms of the matter-impulse-stress tensor T and then claiming that $T = 0$ and $g \neq 0$ represent empty space-time points is not convincing.[11] In general, g represents the gravitational energy and the so-called vacuum solutions exist only in the real world as approximations where the source expressions are ignored. In addition, GTR is not a theory of matter, rather it is an abstraction from matter, and the introduction of a theory of matter via quantum theory gives vacuum solutions different from zero.

If space and time take part in specifying the identity conditions of concrete particulars, then space-time itself cannot be a concrete particular. My suggestion is that it is an abstract particular in the sense that it is existentially dependent on fields and matter. Earman and others reach the substantivalist position by hypostatizing space-time points as objects which are then thought of as the subject for predication of the field properties.[12] Here it seems as if Earman merely hypostatizes the diverse conceptual levels of differential geometry. We begin didactical-mathematically with a differential manifold, then we supply it with diverse affine, metric, and topological structures, and without any further argument it is taken for granted that this pure manifold exists ontologically independently of the structural features which characterize the actual world. What is problematic in the first place is the very idea that we are allowed to hypostatize space-time points as independent entities with their own criteria of identity.

In a recent paper Oliver Pooley (2006) takes issue with Earman and Norton's hole argument. Following Belot and Earman (2001: 228), he defines *sophisticated substantivalism* as any position that denies haecceitistic differences.[13] Such a position regards two diffeomorphic models as representations of the same possible world so they are not injured by the hole argument. In contrast to Belot and Earman, Pooley holds the view that as a sophisticated substantivalist one can argue that space-time points are real substances although their numerical distinctness is grounded by their position in a structure. He believes that such a modest structuralist position does not "go beyond an acceptance of the 'purely structural' properties of the entities in question," while at the same time maintaining that these objects cannot be reduced to the properties and relations themselves. I wonder, however, how space-time points, in terms of their mathematical structure, become physical space-time points. Pooley does not provide us with one single argument according to which the numerical distinctness of the mathematical objects of a manifold (points), whose identity, I agree, depends on their positions in a mathematical structure, corresponds to the numerical distinctness of *real* physical space-time points.

gravitational field and as such "The metric/gravitational field has acquired most, if not all, the attributes that have characterized matter (as opposed to spacetime) from Descartes to Feynman." In contrast, the non-reductive relationist would say the actual geometry is an exemplification of infinitely many possible geometries and that physical space-time seems to gain individuality by being *instantiated* by the gravitational field.

11 Friedmann (1983), p. 223.

12 See for instance Earman (1989), p. 155.

13 See also Belot and Earman (1999).

Let me illustrate why I think Pooley's suggestion that space-time points are real entities in spite of their purely structural properties is problematic. Take a series of identical billiard balls and add an ordering structure: 1, 2, 3, 4, 5,..., from the left, then the identity qua 'number 4 from the left' is given in virtue of the entire structure, namely all the other billiard balls plus the given structure. But it is not a property of that particular ball that if we change it and the 5th ball around, then their identity switches too. They keep their own identity before and after the switching although the order itself is completely unaltered. Space-time points, however, are only defined in relational terms, which mean that they change identity whenever they change their place in the structure. Had they not changed identity, and were they still individuated only by their place in the structure, then the order amongst themselves would not have stayed the same. Analogously, the identity of the number 4 is defined by its place in the entire sequence of numbers, and whatever whole number that may occupy the place between 3 and 5 would be identical with number 4. Here numbers and space-time points seem to be ontologically on par.

Elsewhere I have argued that the structural realism holds an indefensible position on the relationship between mathematically formulated models and the world, namely that there exists an isomorphic coherence between the mathematical structures, which exist independently of the world, and the real structure of the world as it exists independently of mathematics. It does not suffice for the structural realist to point to the ontological commitments of structures given to us by theories (Faye, 2006a). The commitment to a certain structure is always internal to the mathematical framework. The structural realist needs to point to some external commitments which guarantee the existence of real physical counterparts. Claiming conversely that the identity of physical space-time points constitutes a primitive fact which does not require any further explanation is to my mind based on an act of fiat.

Seeing space-time points as independently existing entities with their own identity conditions seems to be a problematic extrapolation of common sense ontology according to which physical objects have intrinsic and not only relational properties. Space-time points lack intrinsic features and without these they do not have a physical basis for differentiation amongst themselves. What determines the identity of space-time points as abstract objects is the mathematical structure as a whole in the sense that we define and identify the constituents (points etc.) within the entire structure; that is to say, the identity of any particular constituent is given in virtue of all the other constituents plus a certain order among them. However, we cannot define and identify the entire structure in virtue of the structure itself. There is therefore a categorical difference between the constituents of the structure (points) and the structure as a whole. Their criteria of identity are not the same. We identify and define the constituents (points) within the structure, but such an individuation is not possible with respect to the structure itself.

So far I have argued that space-time points and fields are ontologically distinct entities because they belong to different kinds of existence, but I have also claimed that space-time is ontologically as well as conceptually posterior to changing and extending things. I now want to conclude that space and time, or space-time, is

nothing but an ordered *set* of all concrete particulars. We need space and time, or space-time, to assist us in identifying and ordering things and events. We would be unable to identify particular things as such and track them down unless we had the possibility of referring to their continuity through space and time. Space gives us the conceptual tool to describe movement of the same material object, and time gives us the conceptual tool to talk about the persistence of numerically the same object possessing different properties. We may therefore say that *particular* concrete objects are in space and time, meaning that they are parts of the ordered set of all changing things. Thus, a persisting object is one that may undergo changes in time, while it continues to stay the same during its changes. The question then arises: is space-time a mere conceptual tool, an instrument for predication, or does it have some sort of reality?

I am inclined to hold that space-time exists as an abstract particular, i.e., as a non-concrete entity, in the sense that its existence is ontologically, but not causally, dependent on the existence of changing things and extended objects. I agree with Jonathan Lowe that the notion of a set is precisely the notion of a number of things and not a 'collection' of things, where natural numbers are kinds of sets (Lowe, 1998: 220–21). Hence space-time is the total number of events and things which exist in the universe from the beginning to the end.

The view I advocate is: space-time exists as an ordered set of all changing things because all its members exist, the set constitutes space-time. Our conception of the set is acquired by acquaintance with a limited number of members of the set and the order is subsequently abstracted from their relations to all possible members. But from this it does not follow that the term used for that abstracted concept refers to an abstract object over and above the entire collection of concrete members in the universe. However, being an ordered set space-time exists as an abstract entity with its own internal identity conditions, and therefore space-time is not reducible to a mere collection. As an abstract entity space-time has no space-bounded or time-bounded properties, it is subject to only tenselessly true predication as far as its relational properties are concerned.

Thus, space-time is not only a set but an ordered set of all concrete particulars in the universe. Any particular event may coexist with some particular events, or precede or succeed some other particular events, and based on these facts we may assign an order of simultaneousness, as well as being earlier or later to all the space-time points which represent the events. In general, events causally (and perceptually) succeed each other and therefore belong to different subsets (hyperplanes) of coexisting events. The actual spatio-temporal order supervenes on structural features of concrete thing and events. What kind of supervenience relation we are talking about has to be dealt with elsewhere due to the lack of space. Grounding the order of space-time points in the causal structure of some actual events we are able to ascribe a unique and unambiguous order to all events in the universe. Every space-time point is ordered with respect to every other space-time point, and we may use this abstracted representation to order any particular event. Indeed, space-time is an abstract entity which has a very privileged relation to the physical reality. There are many mathematical geometries all of which could represent the actual world, but which of the mathematical structures that is instan-

tiated as the actual space-time depends on the distribution of fields and matter in the universe.

ACKNOWLEDGEMENT

I wish to thank Jens Hebor, Rögnvaldur Ingthorsson, Mauro Dorato, and an anonymous referee for their critical comments and helpful suggestions.

REFERENCES

Alexander, H.G., 1956. The Leibniz–Clarke Correspondence. Manchester University Press, Manchester.

Aristoteles, 1955. Physics, W.D. Ross-Edition. Clarendon Press, Oxford.

Belot, G., Earman, J., 1999. From metaphysics to physics. In: Butterfield, J., Pagonis, C. (Eds.), From Physics to Philosophy. Cambridge University Press, Cambridge, pp. 166–186.

Belot, G., Earman, J., 2001. Pre-Socratic quantum gravity. In: Callender, C., Huggett, N. (Eds.), Physics Meet Philosophy on the Planck Scale. Cambridge University Press, Cambridge, pp. 213–255.

Butterfield, J., 1989. Albert Einstein meets David Lewis. In: Fine, A., Leplin, J. (Eds.), In: PSA 1988, vol. 2. Philosophy of Sciences Association, East Lansing, MI.

Earman, J., 1989. World Enough and Space-Time. MIT Press, Boston.

Earman, J., Norton, J.D., 1987. What price substantivalism? The Hole story. British Journal for the Philosophy of Science 38, 515–525.

Einstein, A., 1955. The Meaning of Relativity. Princeton University Press, Princeton.

Faye, J., 1989. The Reality of the Future. Odense University Press, Odense.

Faye, J., 2006a. Science and reality. In: Andersen, H.B., Christiansen, F.V., Hendricks, V., Jørgensen, K.F. (Eds.), The Way through Science and Philosophy: Essays in honour of Stig Andur Pedersen. College Publications, London, pp. 137–170.

Faye, J., 2006b. Is time an abstract entity? In: Stadler, F., Stöltzner, M. (Eds.), Time and History (Series), Proceedings of the 28 International Ludwig Wittgenstein Symposium 2005. Ontos Verlag, Frankfurt, pp. 85–100.

Friedmann, M., 1983. Foundation of Space-Time Theories. Princeton University Press, Princeton.

Gassendi, P., 1972. Selected Works of Pierre Gassendi. Johnson Reprint Corporation, New York. Edited and translated by Craig B. Brush.

Hall, A.R., Hall, M.B. (Eds.), 1962. Unpublished Scientific Papers of Isaac Newton. Cambridge University Press, Cambridge.

Hoefer, C., 2000. Energy conservation in GTR. Studies in History and Philosophy of Modern Physics 31 (2), 187–199.

Lowe, E.J., 1998. The Possibility of Metaphysics. Substance, Identity, and Time. Clarendon Press, Oxford.

Maudlin, T., 1989. The essence of spacetime. In: Fine, A., Leplin, J. (Eds.), In: PSA 1988, vol. 2. Philosophy of Sciences Association, East Lansing, MI.

Maudlin, T., 1990. Substances and spacetime. What Aristotle would have said to Einstein. Studies in History and Philosophy of Science 21, 531–561.

Norton, J.D., 1985. What was Einstein's principle of equivalence? Studies in History and Philosophy of Science 16, 203–246.

Norton, J.D., 1988. The hole argument. In: Fine, A., Leplin, J. (Eds.), In: PSA 1988, vol. 2. Philosophy of Sciences Association, East Lansing, MI, pp. 56–64.

Norton, J.D., 2004. The hole argument. In: Stanford Encyclopedia of Philosophy. CSLI, Stanford University. http://plato.stanford.edu.

Pooley, O., 2006. Points, particles, and structural realism. In: Rickles, D., French, S., Staatsi, J.T. (Eds.), The Structural Foundations of Quantum Gravity. Oxford University Press.
Rovelli, C., 1997. Halfway through the Woods: Contemporary research on space and time. In: Earman, J., Norton, J. (Eds.), The Cosmos of Science. University of Pittsburgh Press, pp. 180–223.
Sklar, L., 1974. Space, Time and Spacetime. University of California Press, Berkeley.

CHAPTER 4

Persistence and Multilocation in Spacetime

Yuri Balashov*

Abstract

The chapter attempts to make the distinctions among the three modes of persistence—endurance, perdurance and exdurance—precise, starting with a limited set of notions. I begin by situating the distinctions in a generic spacetime framework. This requires, among other things, replacement of classical notions such as 'temporal part', 'spatial part', 'moment of time' and the like with their more appropriate spacetime counterparts. I then adapt the general definitions to Galilean and Minkowski spacetime and consider some illustrations. Finally, I respond to an objection to the way in which my generic spacetime framework is applied to the case of Minkowski spacetime.

1. INTRODUCTION. ENDURING, PERDURING AND EXDURING OBJECTS IN SPACETIME

How do physical objects—atoms and molecules, tables and chairs, cats and amoebas, and human persons—persist through time and survive change? This question is presently a hot issue on the metaphysical market. Things were very different some forty years ago, when most philosophers did not recognize the question as an interesting one to ask. And when they did, the issue would quickly get boiled down to some combination of older themes. Here is a cat, and there it is again. It changed in-between (from being calm to being agitated, say); but what is the big deal? Things change all the time without becoming distinct from themselves (as long as they do not lose any of their essential properties, some would add). What else is there to say?

Today we know that there is much more to say. The problem of persistence has become, in the first place, a problem in *mereology*, a general theory of parts

* University of Georgia, Athens, USA

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00004-1

© Elsevier BV
All rights reserved

and wholes.[1] It has also become an issue in a theory of *location*.[2] These two topics continue to drive the debate, especially when it comes to situating the rival accounts of persistence in the "eternalist" spacetime framework. There is a sense in which enduring objects are three-dimensional and multilocated in spacetime whereas perduring objects are four-dimensional and singly located. They are extended in space and time and have both spatial and temporal parts.[3] The latter is strictly denied by endurantists.[4] It is also clear that in view of multilocation in spacetime, the possession of momentary properties and spatial parts by enduring objects must be relativized to time, one way or another.[5] Even in the absence of precise definitions of 'endurance' and 'perdurance',[6] the contrast between these views is very clear. Indeed, the contrast shows up in the labels which are often used to refer to these views: 'three-dimensionalism' and 'four-dimensionalism'.

For quite some time four-dimensionalism had been taken to entail perdurantism, the doctrine that ordinary continuants (rocks, tables, cats, and persons) are temporally extended and persist over time much like roads and rivers persist through space. Recently, however, a different variety of 'four-dimensionalism' has emerged as a leading contender in the persistence debate. According to *stage* theory, ordinary continuants are instantaneous stages rather than temporally extended perduring "worms". Such entities persist by *exduring* (the term due to Haslanger (2003))—by having temporal counterparts at different moments. The distinction between perdurance and exdurance is evident (even though the misleading umbrella title 'four-dimensionalism' gets in the way): perduring and exduring objects have different numbers of dimensions (assuming that exduring object stages are temporally unextended).

On the other hand, the distinction between endurance and exdurance is less clear. Exduring objects lack temporal extension, are three-dimensional, and there is a sense in which they are wholly present at multiple instants. But the same is true of enduring objects. Indeed, the features just mentioned—the lack of temporal extension and multilocation in spacetime—are widely believed to be the distinguishing marks of endurance. How then is exdurance different from endurance?

To be sure, there is a sense in which an exduring object is *not* multiply located. But this is not a sense that can be adopted by someone who wants to regard exdurance as a species of *persistence*, for on that sense, exduring objects do not persist.

[1] For an authoritative exposition of classical mereology, see Simons (1987).

[2] For an authoritative and systematic treatment of theories of spatial location, see Casati and Varzi (1999).

[3] Persisting by being singly or multilocated in spacetime and persisting by having or lacking temporal parts are, arguably, two distinct issues. The distinction is made clear by the conceptual possibility of temporally extended simples (Parsons, 2000) and instantaneous statues (Sider, 2001: 64–65). For the most part I abstract from such possibilities in what follows (but see note 27). For a detailed discussion of the two issues and the resulting four-fold classification of the views of persistence, see Gilmore (2006).

[4] Except in certain exotic cases, such as those briefly considered at the end of Section 2.

[5] Ways in which this can be done have been discussed, among many others, by Lewis (1986: 202–204), Rea (1998), Hudson (2001, 2006), Sider (2001), Hawley (2001) and Haslanger (2003). I revisit the issue in Sections 3 and 4.

[6] Much effort has gone recently into defining 'endurance' and 'perdurance', as well as the underlying notion of *being wholly present* at a time. See, e.g., Merricks (1999), Sider (2001: Chapter 3), Hawley (2001: Chapters 1 and 2), McKinnon (2002), Crisp and Smith (2005), Gilmore (2006), Sattig (2006), and references therein. Some authors are skeptical of the prospect of providing fully satisfactory such definitions that would be acceptable to all parties. See, especially, Sider (2001: 63–68). For a recent attempt to define 'wholly present' in a universally acceptable way, see Crisp and Smith (2005). Even if perfect definitions are not forthcoming, all parties agree that the views in question are transparent enough to debate their merits.

Something persists only if it exists at more than one moment,[7] and an instantaneous object stage, strictly speaking, does not. One could, of course, choose to accept this consequence and agree that exduring objects do not persist. That, however, would undermine the claim of the advocates of stage theory that theirs is the best unified account of *persistence*.[8]

The friends of this account should therefore be sufficiently broad about ways in which an object can be said to be wholly present, or located, at a time. The sense in which this is true of an exduring object is similar to the sense in which an object such as David Lewis is present, located or exists at multiple possible worlds of modal realism. Lewis can be said to exist at world w just in case he has a (modal) counterpart in that world. Similarly, an exduring object can be said to be located (in the requisite broad sense) at t just in case it has a (non-modal) counterpart located (in the strict and narrow sense) at t. This is the only sense in which an exduring object can be said to persist. But as just indicated, on that sense, exduring objects are located at multiple times and share this property with enduring objects. This raises the problem of defining exdurance as a mode of persistence that is different from endurance, as well as perdurance.

Below I attempt to make the distinctions among the three modes of persistence more precise, starting with a limited set of notions. I begin (Section 2) by situating the distinctions in a generic spacetime framework.[9] This requires, among other things, replacement of classical notions such as 'temporal part', 'spatial part', 'moment of time' and the like with their more appropriate spacetime counterparts. I then adapt the general definitions to Galilean (Section 3) and Minkowski (Section 4) spacetime (which is my real target) and consider some illustrations. In Section 5 I focus on an objection to the way in which the generic spacetime strategy of Section 2 is applied to the case of Minkowski spacetime (in Section 4). The objection is due to Ian Gibson and Oliver Pooley (2006) and raises some broader issues of philosophical methodology, which are also discussed in Section 5.[10]

2. PERSISTENCE AND MULTILOCATION IN GENERIC SPACETIME

The task of this section is to develop a framework for describing various modes of persistence in spacetime that would be sufficiently broad to accommodate classical as well as relativistic structures. This requires generalizing some notions that

[7] This is widely accepted as a necessary condition of persistence. The locus classicus is probably Lewis (1986: 202): "Something *persists* iff, somehow or other, it exists at various times."

[8] See Sider (2001: 188–208), Hawley (2001: Chapters 2 and 6), and Varzi (2003).

[9] The generic spacetime approach of Section 2 has much in common with the strategies developed in Rea (1998), Balashov (2000b), Sider (2001: 79–87), Hudson (2001, 2006), Gilmore (2004, 2006), Crisp and Smith (2005), and Sattig (2006). Some of my terminology and basic notions come from Gilmore (2004). Some of the material of Section 2 is based, with modifications and corrections, on an earlier short note (Balashov, 2007) published in *Philosophical Studies* and is used here with kind permission of Springer Science and Business Media. After the publication of Balashov (2007) (and when a draft of the present chapter was finished) I became aware of Thomas Bittner and Maureen Donnelly's paper (Bittner and Donnelly, 2004), which develops a rigorous axiomatic approach to explicating the mereological and locational notions central to the debate about persistence. The approach is set in a broadly classical context but could, I think, be usefully extended to the genetic spacetime framework.

[10] Gibson and Pooley use their objection as a springboard for a sustained criticism of my older arguments (Balashov, 1999, 2000a) in favor of a particular view of persistence (viz., perdurance) over its rivals (i.e., endurance and exdurance) in the context of special relativity (Gibson and Pooley, 2006, Sections 3 and 6). My response to that criticism will have to await another occasion.

figure centrally in the debate about persistence and, as a prerequisite, introducing some underlying spacetime concepts.

2.1. *Absolute chronological precedence*. We shall take the relation of *absolute chronological precedence* ($\prec$) as undefined. Informally, spacetime point p_1 stands in this relation to p_2 ($p_1 \prec p_2$) just in case p_1 is earlier than p_2 in every (inertial) reference frame.[11] It is natural to assume that absolute chronological precedence is asymmetrical ($p_1 \prec p_2 \rightarrow \neg p_2 \prec p_1$) and, hence, irreflexive ($\neg p \prec p$).

2.2. *Achronal regions*. Next we define the notion of an *achronal spacetime region*. A spacetime region (i.e., a set of spacetime points) is achronal iff no point in it absolute-chronologically precedes any other point.

(D1) Spacetime region R is *achronal* $=_{\text{df}} \forall p_1, p_2\ (p_1, p_2 \in R \rightarrow \neg p_1 \prec p_2)$.

Achronal regions are three-dimensional "slices" through spacetime that generalize the classical notion of a moment of time. In fact, a moment of time could be defined as a *maximal* achronal region of spacetime with a certain property:

(D2) R is a *moment of time* $=_{\text{df}}$
[(i) R is a maximal achronal region of spacetime; (ii) R is Ω] $=_{\text{df}}$
$[(\forall p_1, p_2)[p_1, p_2 \in R \rightarrow \neg p_1 \prec p_2] \wedge (\forall p)((\forall p_1, p_2)[p_1, p_2 \in R \cup \{p\} \rightarrow \neg p_1 \prec p_2] \rightarrow p \in R) \wedge R$ is $\Omega]$.

Clause (ii) is needed because nothing in the above definition requires an achronal region to be a *flat* 3D hypersurface in spacetime. But it is natural to suppose that no achronal hypersurface can represent a moment of time in the classical or special relativistic setting unless it is flat. In these settings, 'Ω' could be taken to be synonymous with 'flat', where flatness is defined in the usual metric way.[12]

The significance of flat achronal hypersurfaces in special relativistic spacetime and their relation to the notion of time are issues that require more discussion and I shall return to them in Section 5. But they do not play any part in the general definitions of the different modes of persistence provided later in this section. What does play a central role in them is the notion of achronality (and the underlying relation of absolute chronological precedence). My approach takes the second notion as a starting point to allow maximum generality. But in familiar contexts, it bears close relationship to other widely used concepts. Thus in many applications, a maximal achronal region is none other than a Cauchy surface—a spacelike hypersurface that intersects every unbounded timelike curve at exactly one point. But there is no need to invoke additional notions, such as 'spacelike' and 'timelike', in a generic context where all the useful work could be done by 'achronal'.

We need, however, make a brief digression to note a familiar problem with the concepts of 'absolute chronological precedence' and 'achronal', which is brought

[11] As one would expect (see Sections 3 and 4 below), in classical spacetime absolute chronological precedence can be taken literally to mean precedence in the absolute time while in the special relativistic framework absolute chronological precedence is equivalent to the frame-invariant relation in which two points stand just in case they are either (i) timelike separated or (ii) lightlike (null) separated while being distinct.

[12] What about general relativity? Although it goes beyond the scope of this chapter, it is worth noting that, except in very special cases (e.g., certain idealized cosmological models), the notion of a moment of time lacks any meaning in general relativistic spacetime.

to light by considering peculiar spacetimes possessing closed or "almost closed" timelike curves. For the purpose of this informal consideration, 'timelike' could be taken to be synonymous with "non-achronal". Closed timelike curves exist, for example, in Gödelian cosmological models of general relativity, but a flat "cylindrical" spacetime could serve as a useful toy model.[13] It is easy to see that there is a sense in which two "nearby" points p_1 and p_2 can stand in the relation of absolute chronological precedence ($p_1 \prec p_2$)—the sense obtained by tracing a non-achronal curve from p_1 to p_2 around the "cylinder". But there is also a sense in which they are not ($\neg p_1 \prec p_2$)—the sense obtained by tracing an achronal curve from p_1 to p_2 along a generatrix of the "cylinder". Accordingly, a certain region containing both p_1 and p_2 might be classified by (D1) as being both achronal and non-achronal. Situations of this sort figure prominently in the literature on time travel.[14]

Another problem arises in spacetimes having a "trouser" topology.[15] Points p_1 and p_2 belonging to different legs of the "trousers" do not bear any well-defined metrical relations to each other and, hence, are not related by $\prec$. But if p_1 precedes the merger by just a few seconds but p_2 is thousands of years away from it, there is some inclination to say that p_2 chronologically precedes p_1 (in the sense associated with "$\prec$").

Both problems could perhaps be alleviated by making the definition of 'achronal region' in the relevant sense *local*[16] and thus consistent with closed or "almost closed" timelike curves, and with the "trouser" topology. We shall abstract from such situations in what follows. This limitation is quite tangential to the main task of the chapter—to capture the distinctive features of the various modes of persistence in a spacetime setting by using an economical set of primitive notions. We shall assume, accordingly, that global maximal achronal regions are always available.

2.3. *LOCATION*. Persisting objects are located at regions of spacetime. For our purposes, 'located at' means *exactly located*. The guiding idea here is that the region at which an object is exactly located is the region into which the object exactly fits and which has exactly the same size, shape, and position as the object itself.[17]

I take 'located at R' to mean the same as 'wholly present at R', but I put aside the question of whether the latter notion can be rigorously defined for objects having (achronal) parts.[18] Providing such a definition is one of the most intensely debated problems nowadays.[19] My concerns here are, however, rather orthogonal to it, for I am interested in the underlying sense of 'located at R' applicable to (achronally) composite and non-composite objects alike, which any such definition must take as a starting point.

[13] Cf. Gilmore (2007), where a similar toy model is used to investigate the implications of time travel scenarios for the issue of persistence.

[14] For recent discussions, see Gilmore (2006, 2007) and Gibson and Pooley (2006: Section 5).

[15] See Gilmore (2004: 204, notes 19 and 20), who refers in this connection to Sklar (1974: 306–307).

[16] See Gilmore (2006: 209, note 19) for one attempt to do it.

[17] This notion of exact location is similar to Gilmore's notion of *occupation* (Gilmore, 2006), Hudson's notion of *exact occupation* (Hudson, 2001), Bittner and Donnelly's notion of *exact location* (Bittner and Donnelly, 2004), and other equivalents found in the recent literature. But see Parsons (2007) for a very different notion of 'exactly located'.

[18] See the definition of *achronal part* below (D6).

[19] See Rea (1998), Sider (2001: 63–68), McKinnon (2002), Crisp and Smith (2005), Parsons (2007) and references therein.

What is essential to my task is that there be a common notion of location—call it 'LOCATION'—which is broad enough to incorporate the modes in which both enduring and exduring objects are capable of multilocation. To repeat, the sense in which an exduring object accomplishes this feat is similar to the sense in which a worldbound individual of the Lewisian pluriverse can nonetheless be said to exist at multiple worlds. To make the notion of LOCATION precise, let us start with (non-modal) counterparthood and stipulate that every object (enduring, perduring, or exduring) is a (non-modal) counterpart of itself. This is a natural assumption that does not impose any undue commitments on endurantism or perdurantism. The advocates of both theories could agree that every persisting object has an "improper" non-modal counterpart: itself—multiply located in the case of endurantism, and singly located in the case of perdurantism.[20] 'LOCATED at *R*' could then be defined as follows:

(D3) *o* is (exactly) *LOCATED* at $R =_{df}$ one of *o*'s (non-modal) counterparts is (exactly) located at *R*.

The following definitions (adapted from Gilmore (2004: Chapter 2) and (2006: 204ff)) help to align LOCATION more precisely with the notion of persistence.[21]

(D4) Spacetime region ***o*** is the *path* of object $o =_{df}$ ***o*** is the union of the spacetime region or regions at which *o* is LOCATED.

(D5) *o persists* $=_{df}$ *o*'s path is non-achronal.

The advantage of (D1) and (D3)–(D5) lies in their ability to offer a *unified* account of persistence and multilocation, on which (i) enduring, perduring and exduring objects persist in the same sense, and (ii) enduring and exduring objects are multilocated in the same sense. All parties can agree that endurance, perdurance, and exdurance are *bona fide* modes of persistence and, in particular, that exdurance is not a second-class citizen: exduring objects persist in the same robust sense as enduring objects do. This allows one to focus on the important question of how they manage to do so.

2.4. *Achronal and diachronic parts*. Next we need generalizations of the concepts of spatial and temporal part. We shall take a three-place relation '*p* is a part of *o* at achronal region *R*'—as a primitive.[22] The intuitive ancestor of this relation is the familiar time-relativized sense in which certain cells are part of me at one time but not at another. Where *p*, *o* and *R* stand in this relation, we shall say that *p* is an *achronal part* of *o* at achronal region *R* and denote it with the subscript '$\perp$':

[20] For those who may be inclined to resist this usage of 'counterpart' as too stretched, a somewhat less elegant equivalent of (D3) is readily available:

(D3′) *o* is (exactly) *LOCATED* at $R =_{df}$ *o* is exactly located at *R* or one of *o*'s (non-modal) counterparts is (exactly) located at *R*.

[21] But Gilmore might object to combining his definition of 'persists' with the broad sense of 'LOCATED at'. On his official view, as far as I can see, exduring objects do not persist. See, however, Gilmore (2006: 230, note 21), where he suggests a rather innocuous modification to his approach that would accommodate exdurance.

[22] This relation is similar to that used by Hudson (2001) in developing his Partist view of persistence but more restrictive than the latter (and thus closer to the familiar concept of temporary parthood), in that Hudson's notion relativizes parthood to arbitrary regions of spacetime whereas mine is limited to achronal regions.

(D6) $p_\perp$ is an *achronal part* of o at achronal spacetime region $R =_{\text{df}} p_\perp$ is a part of o at R.

Diachronic parthood could then be defined as follows:

(D7) $p_\parallel$ is a *diachronic part* of o at achronal spacetime region $R =_{\text{df}}$ (i) $p_\parallel$ is located at R but only at R, (ii) $p_\parallel$ is a part of o at R, and (iii) $p_\parallel$ overlaps at R everything that is a part of o at R.

Note that neither p nor o need be "as large as" the achronal region R, in order to stand in the relation 'p is a part of o at R'. All that could reasonably be required of the achronal extents of o and p at R, is that the intersection of p's path with R be "within" the intersection of o's path with R:

(WITHIN) p is a part of o at achronal region $R \rightarrow \boldsymbol{p} \cap R \subseteq \boldsymbol{o} \cap R$.

This, of course, entails that both $\boldsymbol{p} \cap R$ and $\boldsymbol{o} \cap R$ are "within" R. Thus my hand is a part of me at a certain momentary location of my hand, at a momentary location of my body, and at a momentary location of the Solar system. Furthermore, if I am an exduring object my hand is a part of me at an achronal region at which neither I nor my hand are even "sub-located"—say, a region at which I *was* located at some moment 10 years ago. In this case the job of grounding R-relativized parthood is done by the non-modal counterparts of the relevant objects. Finally, assuming perdurance, one of my cells at t (i.e., a global moment of time) is a part of me at my momentary location at t (i.e., at the location of my momentary t-part), but also a part of me at the momentary location of the Solar system at t.[23]

In contrast, the notion of diachronic parthood is more restrictive: if $p_\parallel$ is a diachronic part of o at achronal region R then $p_\parallel$ must "fit into" R exactly, although o may "overfill" R in virtue of having parts (both achronal and diachronic) at superregions of R.

In the subsequent discussion the generic relations of achronal and diachronic parthood, explicated in (D6) and (D7), are restricted to distinguished achronal regions—those "containing" (in a relevant sense) the objects involved in the relation. Such regions are *achronal slices* of the objects' paths.

(D8) $R_\perp$ is an *achronal slice* of $R =_{\text{df}} R_\perp$ is a non-empty intersection of a maximal achronal 3D region with $R =_{\text{df}} (\exists R^*)[(\forall p_1, p_2)(p_1, p_2 \in R^* \rightarrow \neg p_1 \prec p_2) \wedge (\forall p)((\forall p_1, p_2)[p_1, p_2 \in R^* \cup \{p\} \rightarrow \neg p_1 \prec p_2] \rightarrow p \in R^*) \wedge R_\perp = R \cap R^* \wedge (\exists p) p \in R_\perp]$

More comments are in order.

(i) As defined by (D6) and (D7), achronal and diachronic parthood are not mutually exclusive. Indeed, diachronic parthood is just a special case of achronal parthood. In the case of both perdurance and exdurance, the diachronic part of any object at a t-slice of its path is equally its achronal part at that slice.

[23] One counterintuitive consequence of R-relativized parthood thus understood must be noted: p may be a part of o at an achronal region "not large enough" for o, provided that it is "large enough" for p. For example, (WITHIN), as stated above, does not preclude me from being a part of my hand at a momentary location of my hand. A fully axiomatic treatment of R-relativized parthood would probably need to rule out such cases, perhaps by modifying (WITHIN). This would lead to complications that are best avoided in the present context.

(ii) However, there is a sense in which *proper* achronal and diachronic parthood are exclusive. If *proper parthood at achronal region R* is defined as *asymmetrical* achronal parthood at *R*:

(D9) $p_\perp$ is a *proper achronal part* of o at achronal region $R =_{df}$ (i) $p_\perp$ is an achronal part of o at R, (ii) o is not an achronal part of $p_\perp$ at R,

then, if $p_\perp$ is a proper achronal part of o at some achronal slice $\boldsymbol{o}_\perp$ of its path then $p_\perp$ is not a diachronic part of o at $\boldsymbol{o}_\perp$, proper or not. The reason, roughly, is that $p_\perp$ is "smaller" than o at $\boldsymbol{o}_\perp$ and thus cannot be a diachronic part of o at $\boldsymbol{o}_\perp$.

And if *proper diachronic parthood* is defined as *asymmetrical* diachronic parthood:

(D10) $p_\parallel$ is a *proper diachronic part* of o at achronal region $R =_{df}$ (i) $p_\parallel$ is a diachronic part of o at R, (ii) o is not a diachronic part of $p_\parallel$ at R,

then, if $p_\parallel$ is a proper diachronic part of o at some achronal slice $\boldsymbol{o}_\perp$ of its path then $p_\parallel$ is not a proper achronal part of o at $\boldsymbol{o}_\perp$. The reason, roughly, is that being a diachronic part of o at $\boldsymbol{o}_\perp$, proper or not, makes $p_\parallel$ "as large as" o at $\boldsymbol{o}_\perp$ and, hence, not a proper achronal part of it at $\boldsymbol{o}_\perp$. However, $p_\parallel$ and o will in general be *improper* achronal parts of each other at $\boldsymbol{o}_\perp$.

On the other hand, if proper achronal and diachronic parthood at achronal region R are understood as follows:

(D9′) $p_\perp$ is a *proper achronal part* of o at achronal region $R =_{df}$ (i) $p_\perp$ is an achronal part of o at R, (ii) $p_\perp \neq o$;

(D10′) $p_\parallel$ is a *proper diachronic part* of o at achronal region $R =_{df}$ (i) $p_\parallel$ is an diachronic part of o at R, (ii) $p_\parallel \neq o$,

then one object could be both a proper achronal and a proper diachronic part of another object at some achronal region. Consider a perduring or exduring statue and the piece of clay of which it is composed. Some would argue that the statue (and, hence its t-part) is not identical with the piece of clay (and its corresponding t-part). If so then the statue and the piece of clay are both proper achronal and proper diachronic parts of each other at the t-slice of the path of both objects.

(D9), (D10), (D9′) and (D10′) raise an interesting question of how to develop general R-relativized mereology.

(iii) As defined, achronal and diachronic parts are achronal, that is, diachronically (or "temporally", where this designation is appropriate) non-extended. In this I deviate from the authors who explicitly allow temporally extended temporal parts and make them do some useful work.[24]

2.5. *Achronal Universalism*. Finally, we assume the thesis of Achronal Universalism:

(Achronal Universalism)

(i) Any enduring object is located at every achronal slice of its path;
(ii) any perduring object has a diachronic part at every achronal slice of its path;
(iii) any exduring object is LOCATED at every achronal slice of its path.

[24] See, in particular, Heller (1990), Zimmerman (1996), Butterfield (2006) and note 28 below.

Thus in the context of this general consideration, which is not specific to any particular type of spacetime, we impose no restriction whatsoever on which of the achronal slices of an object's path contain that object or one of its diachronic parts. Nothing of substance turns on this simplifying assumption for the purpose of this section. The situation will change when we turn to adapting the generic definitions to particular spacetime structures in later sections. At that point, the statement of Achronal Universalism appropriate to a given such structure will become a more controversial matter.

2.6. *'Endurance', 'perdurance' and 'exdurance' defined*. The following definitions capture the important distinctions among the three modes of persistence.

(D11) *o endures* $=_{\text{df}}$ (i) *o* persists, (ii) *o* is located at every achronal slice of its path, (iii) *o* is LOCATED only at achronal slices of its path.

(D12) *o perdures* $=_{\text{df}}$ (i) *o* persists, (ii) *o* is LOCATED only at its path, (iii) the object located at any achronal slice $\boldsymbol{o}_{\perp}$ of *o*'s path is a proper diachronic part of *o* at $\boldsymbol{o}_{\perp}$.

(D13) *o exdures* $=_{\text{df}}$ (i) *o* persists, (ii) *o* is located at exactly one region, which is an achronal slice of its path, (iii) *o* is LOCATED at every achronal slice of its path.

On these definitions, the difference between endurance and perdurance is as expected: (i) enduring but not perduring objects are multilocated (and, hence, multiLOCATED) in spacetime; (ii) perduring but not enduring objects have diachronic parts.[25]

More importantly, the definitions also bring out the crucial distinction between perdurance and exdurance: (a) exduring but not perduring objects are multiLOCATED in spacetime; (b) while both perduring and exduring objects have diachronic parts, perduring objects have only *proper* diachronic parts. That exduring objects have improper diachronic parts follows from clauses (ii) and (iii) of (D13) and the definition of 'diachronic part', which together entail that the object located at every achronal slice of an exduring object's path is a diachronic part, at that slice, of some object: namely, itself.[26]

Finally, the definitions pinpoint the difference between exdurance and endurance: while both exduring and enduring objects are multiLOCATED, only the former (again, barring some exotic cases; see below) have diachronic parts at every

[25] Barring certain exotic exceptions; see note 27.

[26] This does not imply that exduring objects have *only* improper diachronic parts. It depends on how proper diachronic parthood at *R* is defined—the issue already considered above. In any case, the stage theorist should, of course, deny that an exduring object *o* is strictly identical with its t_1-stage, p_1, as well as with its distinct t_2-stage, p_2. If so, then under the aforementioned definition (D10′) of *R*-relativized proper diachronic parthood:

(D10′) $p_{\parallel}$ is a *proper diachronic part* of *o* at achronal region $R =_{\text{df}}$ (i) $p_{\parallel}$ is an diachronic part of *o* at *R*, (ii) $p_{\parallel} \neq o$,

at least one of p_1 and p_2 is a proper diachronic part of *o* (at t_1- or t_2-slice of *o*'s path). On the other hand, if proper parthood at *R* is defined as *asymmetrical parthood* at *R*:

(D10) $p_{\parallel}$ is a *proper diachronic part* of *o* at achronal region $R =_{\text{df}}$ (i) $p_{\parallel}$ is an diachronic part of *o* at *R*, (ii) *o* is not a diachronic part of $p_{\parallel}$ at *R*,

then both p_1 and p_2 are improper parts of *o*, at different *t*-slices of its path. This, of course, does not entail that $p_1 = p_2$.

region at which they are LOCATED. Indeed, clause (ii) of (D11) generally prevents an enduring object from having a diachronic part at any achronal slice of its path.[27]

(D11)–(D13) thus delineate the important contrasts among the three modes of persistence.[28]

3. PERSISTENCE AND MULTILOCATION IN GALILEAN SPACETIME

In this section I adapt the generic framework introduced above to Galilean spacetime. This task is relatively straightforward. The relation of absolute chronological precedence ($\prec$) in Galilean spacetime (ST^G) coincides with the relation of absolute temporal precedence: $p_1 \prec p_2 \leftrightarrow t_1 < t_2$, where (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) are the coordinates of p_1 and p_2 in any Cartesian coordinate system associated with any inertial frame of reference. Accordingly, a region R of Galilean spacetime is achronal iff it is a subregion of an absolute time hyperplane. That is:

(D1^G) Region R of ST^G is *achronal* $=_{df} \forall p_1, p_2\ (p_1, p_2 \in R \rightarrow t_1 = t_2)$.

And a moment of time (= a maximal achronal region) is simply a time hyperplane in ST^G:

(D2^G) R is a *moment of time* in ST^G $=_{df}$ R is a time hyperplane in ST^G.

We take the definitions of *LOCATION* and *path* directly from Section 2.

(D3^G) o is (exactly) *LOCATED* at R in ST^G $=_{df}$ one of o's (non-modal) counterparts is (exactly) located at R.[29]

(D4^G) Spacetime region $\boldsymbol{o}$ is the *path* of object o in ST^G $=_{df}$ $\boldsymbol{o}$ is the union of the spacetime region or regions at which o is LOCATED.

According to our older generic definition (D5), *o persists* just in case o's path is non-achronal. Adapted to Galilean spacetime, this boils down to the requirement that o's path intersect at least two distinct moments of time.

(D5^G) o persists in ST^G $=_{df} \exists p_1, p_2 \in \boldsymbol{o}, t_1 \neq t_2$.

[27] But here (finally!) is an exotic exception. Consider an enduring lump of clay that becomes a statue for only an instant (Sider, 2001: 64–65). On (D7), the statue is a diachronic part of the lump at that instant.

[28] At the same time, it should be emphasized that these definitions are not watertight, and I did not strive to make them so. In fact, one may doubt that watertight definitions are even possible, especially in the case of endurance (see note 6). Apart from Sider's instantaneous statue (note 27), (D11)–(D13) give intuitively wrong results in other exotic cases. Consider an organism composed of perduring cells and stipulate that the cells and their diachronic parts are the *only* proper parts of the organism (Merricks, 1999: 431). By clause (iii) of (D12), the organism itself does not perdure. Another exotic case (suggested by a referee of Balashov (2007)) includes an object satisfying (D11) but having "finitely extended diachronic parts." It is unclear whether such an object could be regarded as enduring. Relatedly, there could be an object satisfying clauses (i) and (ii) of (D12) but having only "finitely extended proper diachronic parts." On (D12), such an object does not perdure, an intuitively wrong result. To handle possibilities of this sort, one would need to make full use of the appropriately defined notion of a "diachronically extended diachronic part," which lies outside the scope of this project. See also note 24. Fortunately, cases of this sort are too remote to bear on the agenda of this chapter and we can safely ignore them. For our purposes, (D11)–(D13) provide good working accounts of the three modes of persistence.

[29] As before, those who are dissatisfied with the broad sense of 'counterpart' at work in (D3^G), may choose a less elegant equivalent of (D3^G):

(D3$^{G\prime}$) o is (exactly) *LOCATED* at region R of ST^G $=_{df}$ o is exactly located at R or one of o's (non-modal) counterparts is (exactly) located at R.

The earlier generic definitions of 'achronal part of o at achronal region R' (D6) and 'diachronic part of o at achronal region R' (D7) generalized the concepts of spatial part and instantaneous temporal part to the spacetime framework. In Galilean spacetime, however, all and only achronal regions are moments of absolute time. This effectively reduces some of the generic notions of Section 2 to their more familiar classical predecessors. In particular, an achronal slice $R_{\perp}$ of R in ST^{G} is simply the intersection of R with a moment of time:

(D8$^{\mathrm{G}}$) $R_{\perp}$ is an *achronal slice* of R in ST^{G} $=_{\mathrm{df}}$ $R_{\perp}$ is a non-empty intersection of a moment of time (i.e., a time hyperplane) with R.

Accordingly, I shall refer to the achronal slice of R at t in ST^{G} simply as 't-slice of R' or '$R_{\perp t}$'. This brings the concepts of achronal and diachronic parthood at achronal region R closer to the older concepts of *temporal part at t* and *spatial part at t*. In what follows I shall sometimes use such simpler notions, where context makes it clear that 't' refers not to an entire hyperplane of absolute simultaneity but to a rather small subregion of it: $\boldsymbol{o}_{\perp t}$.

As before, we assume Achronal Universalism:

(Achronal Universalism$^{\mathrm{G}}$)

(i) Any enduring object is located at every t-slice of its path (in Galilean spacetime);
(ii) any perduring object has a t-part at every t-slice of its path;
(iii) any exduring object is LOCATED at every t-slice of its path.

On this assumption, endurance, perdurance and exdurance in Galilean spacetime can be defined as follows:

(D11$^{\mathrm{G}}$) *o endures* in ST^{G} $=_{\mathrm{df}}$ (i) o persists, (ii) o is located at every t-slice of its path, (iii) o is LOCATED only at t-slices of its path.
(D12$^{\mathrm{G}}$) *o perdures* in ST^{G} $=_{\mathrm{df}}$ (i) o persists, (ii) o is LOCATED only at its path, (iii) the object located at any t-slice of o's path is a proper t-part of o.
(D13$^{\mathrm{G}}$) *o exdures* in ST^{G} $=_{\mathrm{df}}$ (i) o persists, (ii) o is located at exactly one region, which is a t-slice of its path, (iii) o is LOCATED at every t-slice of its path,

As noted in Section 2, these definitions are not watertight, but they bring out all the essential differences among the three modes of persistence in Galilean spacetime.

Multilocation has a familiar consequence for the analysis of temporal predication. Galilean spacetime provides a convenient framework for discussing this issue. To say what properties (and spatial parts) an object has at t, the endurantist who subscribes to spacetime realism must relativize possession of temporary properties (and spatial parts) to time. She cannot say that a certain poker is hot and stop here, because the selfsame poker is also cold, when it is wholly present at a different time.[30] Time must somehow be worked into the picture. One has to explain how time interacts with predication and what makes statements attributing temporary properties to objects true. There are several ways of doing it, which

[30] This is often referred to as the "problem of temporary intrinsics" or the "problem of change."

bring with them somewhat distinct semantics and metaphysics of temporal modification.[31]

In discussions that abstract from spacetime considerations such schemes are often looked upon as providing a semantic regimentation for simple expressions of the form '*o* has Φ at t''. In a more systematic treatment, ascription of properties must be relativized to achronal regions of spacetime, namely, to achronal slices of *o*'s path. However, since in ST^G all the achronal regions of interest can (in ordinary cases) be indexed by moments of absolute time, we can, for the purpose of illustration, keep the simple form.

The following is a brief summary of the analyses of temporal predication in the competing views of persistence, beginning with endurance, which allows three somewhat different schemes:[32]

(EndST^G-1: Rel) Enduring object *o* has Φ at *t* in Galilean spacetime $=_{df}$ *o* bears Φ-at to *t*.
(EndST^G-2: Ind) Enduring object *o* has Φ at *t* in Galilean spacetime $=_{df}$ *o* has Φ-at-*t*.
(EndST^G-3: Adv) Enduring object *o* has Φ at *t* in Galilean spacetime $=_{df}$ *o* has$_t\Phi$.

Perdurance and exdurance, on the other hand, naturally go along with the following canonical accounts of temporal predication in Galilean spacetime:

(PerST^G) Perduring object *o* has Φ at *t* in Galilean spacetime $=_{df}$ *o*'s *t*-part has Φ.
(ExdST^G) Exduring object *o* has Φ at *t* in Galilean spacetime $=_{df}$ *o*'s *t*-counterpart has Φ.

To illustrate these ideas further, consider a 10 meter-long pole in Galilean spacetime. At a certain moment, it starts to contract until its length is reduced to 5 meters. On endurantism, the pole is a 3D entity extended in space but not in time. It is located at all *t*-slices of its path and any such intersection features the full set of properties the pole has at a corresponding time, including its length. Some of these properties are apparently incompatible, such as being 10 meters long and being 5 meters long. How can the self-same object exhibit incompatible properties? Part of the controversy about persistence arises from taking this question seriously. But given multilocation of enduring entities in (Galilean) spacetime, the answers are readily available. On Relationalism, the pole comes to have the property of being 5 meters long at t_1 and 10 meters long at t_2 by bearing the relation *5-meter-long-at* to t_1 and *10-meter-long-at* to t_2.[33] On Indexicalism, the pole accomplishes the same feat by exemplifying two time-indexed properties, *5-meter-long-at-*t_1 and *10-meter-long-at-*t_2. On Adverbialism, the pole possesses the simple property *5-meter-long* in the t_1-ly way, and another property, *10-meter-long*, in a different, t_2-ly way.

[31] The general strategy of relativizing temporary properties to times was sketched by Lewis (1986: 202–204). It was then implemented in a great number of works and in many different forms. For recent contributions and references see MacBride (2001), Haslanger (2003).

[32] In the text below 'Rel' stands for "Relationalism" (not to be confused with spacetime Relationism), 'Ind' for "Indexicalism", and 'Adv' for "Adverbialism". None of these terms is universally accepted, but all are widely used (and sometimes confused with each other) in the literature.

[33] As noted above, in simple contexts 't_1' and 't_1' come in handy as useful shorthand for '$\boldsymbol{o}_{\perp t_1}$' and '$\boldsymbol{o}_{\perp t_2}$'.

On perdurantism, on the other hand, the pole is a 4D entity extended both in space and time. It persists by having distinct momentary t-parts at each t-slice through its path. When we say that the pole is 10 meters long at t_1 and 5 meters long at t_2, what we really mean is that the pole's t_1-part has the former property and its t_2-part the latter. The sense in which the properties of the pole's t-parts can be attributed to the 4D whole is, in many ways, similar to the sense in which the properties of the spatial parts of an extended object are sometimes attributed to the whole. When we say that the oil pipe is hot in the vicinity of the pump and cold elsewhere, we really mean that the pipe has, among its spatial parts, a part in the vicinity of the pump, which is hot, and an elsewhere part, which is cold. Just as the pipe (and entire thing) *changes* from being hot to being cold, the pole (the entire perduring object) changes from being long to being short.

On exdurantism, the pole is a 3D entity LOCATED at multiple t-slices through its path, thanks to having distinct t-counterparts at each such slice. The pole comes to be 10 meters long at t_1 and 5 meters long at t_2 by having a t_1-counterpart and a t_2-counterpart, which have these respective lengths *simpliciter*. (Remember that the t-counterpart relation is reflexive.)

Persistence and temporal predication in Galilean spacetime are straightforward.

4. PERSISTENCE AND MULTILOCATION IN MINKOWSKI SPACETIME

Minkowski spacetime ($\mathrm{ST^M}$) brings novel and interesting features. In $\mathrm{ST^M}$ absolute chronological precedence is the frame-invariant relation in which two points stand just in case they are either timelike separated or lightlike separated while being distinct: $p_1 \prec p_2 \leftrightarrow I(p_1, p_2) \geq 0 \wedge p_1 \neq p_2$, where $I(p_1, p_2) \equiv c^2(t_2 - t_1)^2 - (\mathbf{x}_2 - \mathbf{x}_1)^2$ is the relativistic interval. Accordingly, any spacelike hypersurface[34] counts as an achronal region of $\mathrm{ST^M}$:

(D1$^{\mathrm{M}}$) Region R of $\mathrm{ST^M}$ is *achronal* $=_{\mathrm{df}} \forall p_1, p_2\ (p_1, p_2 \in R \rightarrow I(p_1, p_2) < 0)$.

But only a subset of them—those that are *flat*—represent legitimate perspectives: moments of time in inertial reference frames, $\{t^{\mathrm{F}}\}$:

(D2$^{\mathrm{M}}$) R is a *moment of time* in $\mathrm{ST^M}$ $=_{\mathrm{df}}$ R is a spacelike hyperplane in $\mathrm{ST^M}$.

It is therefore appropriate to index LOCATION of persisting objects and their parts in $\mathrm{ST^M}$ to t^{F}.[35] And it is convenient to treat 't^{F}' as a two-parameter index, assuming that the choice of a particular coordinate system adapted to a given inertial reference frame can somehow be fixed.

Two related facts about frame-relative moments of time in $\mathrm{ST^M}$ are worth noting:

[34] A hypersurface is spacelike just in case any two points on it are spacelike separated.

[35] The appropriateness of restricting "legitimate perspectives" and LOCATIONS of persisting objects and their parts to moments of time in inertial reference frames in $\mathrm{ST^M}$ has been criticized by Gibson and Pooley (2006: 159–165). I discuss and respond to their criticism in the next section.

(i) Any two distinct moments of time t_1^F and t_2^F, $t_1^F \neq t_2^F$, in a single frame F are parallel and, therefore, do not overlap. In this respect, moments of time in a given frame are similar to absolute moments of time in $\mathrm{ST^G}$.

(ii) Any two moments of time in distinct frames, $t_1^{F_1}$ and $t_2^{F_2}$, $F_1 \neq F_2$, overlap. In this respect, moments of time in distinct frames in $\mathrm{ST^M}$ are very different from absolute moments of time in $\mathrm{ST^G}$.

LOCATION and *path* in $\mathrm{ST^M}$ can then be defined.

(D3$^{\mathrm{M}}$) o is (exactly) *LOCATED* at region R of $\mathrm{ST^M}$ $=_{\mathrm{df}}$ one of o's (non-modal) counterparts is (exactly) located at R.[36]

(D4$^{\mathrm{M}}$) Spacetime region $\boldsymbol{o}$ is the *path* of object o in $\mathrm{ST^M}$ $=_{\mathrm{df}}$ $\boldsymbol{o}$ is the union of the spacetime region or regions at which o is LOCATED.

On the generic definition of persistence (D5), *o persists* just in case o's path is non-achronal. In Minkowski spacetime, this is equivalent to the requirement that o's path intersect at least two distinct moments of time in a single frame or, alternatively, that o's path contain two non-spacelike separated points.

(D5$^{\mathrm{M}}$) o persists in $\mathrm{ST^M}$ $=_{\mathrm{df}} \exists p_1, p_2 \in \boldsymbol{o}\ \exists F\ t_1^F \neq t_2^F$

$$=_{\mathrm{df}} \exists p_1, p_2 \in \boldsymbol{o}\ (p_1 \neq p_2 \wedge I(p_1, p_2) \geq 0).$$

As before, $(\mathbf{x}_1^F, t_1^F)$ and $(\mathbf{x}_2^F, t_2^F)$ are the coordinates of p_1 and p_2 in a Cartesian coordinate system adapted to the inertial frame of reference F.

An achronal slice $R_\perp$ of R in $\mathrm{ST^M}$ is the intersection of R with a moment of time in some inertial frame:

(D8$^{\mathrm{M}}$) $R_\perp$ is an *achronal slice* of R in $\mathrm{ST^M}$ $=_{\mathrm{df}}$ $R_\perp$ is a non-empty intersection of a moment of time (i.e., a time hyperplane) with R.

We shall refer to the achronal slice of R at t^F in $\mathrm{ST^M}$ as 't^F-slice of R' or '$R_{\perp t^F}$'. And we shall allow such expressions as 'achronal part of o at t^F', 'diachronic part of o at t^F' and 'o's t^F-part' to go proxy for their more complex equivalents, such as 'achronal part of o at t^F-slice $\boldsymbol{o}_{\perp t^F}$ of o's path $\boldsymbol{o}$' and so forth. Moreover, we shall allow ourselves the liberty to speak of "spatial parts" of persisting objects in Minkowski spacetime when it is clear what reference frame is under consideration.

As before, we adopt a version of Achronal Universalism, appropriate for $\mathrm{ST^M}$:

(Achronal Universalism$^{\mathrm{M}}$)

(i) Any enduring object is located at every t^F-slice of its path (in $\mathrm{ST^M}$);
(ii) any perduring object has a t^F-part at every t^F-slice of its path;
(iii) any exduring object is LOCATED at every t^F-slice of its path.

Given Achronal Universalism$^{\mathrm{M}}$, the definitions of the three basic modes of persistence in Minkowski spacetime are rather simple.

[36] Alternatively:
(D3$^{\mathrm{M}\prime}$) o is (exactly) *LOCATED* at region R of $\mathrm{ST^M}$ $=_{\mathrm{df}}$ o is exactly located at R or one of o's (non-modal) counterparts is (exactly) located at R.

(D11^{M}) *o endures* in ST^{M} $=_{df}$ (i) *o* persists, (ii) *o* is located at every t^F-slice of its path, (iii) *o* is LOCATED only at t^F-slices of its path.

(D12^{M}) *o perdures* in ST^{M} $=_{df}$ (i) *o* persists, (ii) *o* is LOCATED only at its path, (iii) the object located at any t^F-slice of *o*'s path is a proper t^F-part of *o*.

(D13^{M}) *o exdures* in ST^{M} $=_{df}$ (i) *o* persists, (ii) *o* is located at exactly one region, which is a t^F-slice of its path, (iii) *o* is LOCATED at every t^F-slice of its path.

The analysandum of the predication schemes characteristic of endurance, perdurance and exdurance in ST^{M} is an expression of the form '*o* has Φ at t^F' (where, as before, 't^F' is a simplified index for what, in a more systematic treatment, would be '$\boldsymbol{o}_{\perp t^F}$').

(EndST^{M}-1: Rel) Enduring object *o* has Φ at t^F in Minkowski spacetime $=_{df}$ *o* bears Φ-at to t^F.

(EndST^{M}-2: Ind) Enduring object *o* has Φ at t^F in Minkowski spacetime $=_{df}$ *o* has Φ-at-t^F.

(EndST^{M}-3: Avd) Enduring object *o* has Φ at t^F in Minkowski spacetime $=_{df}$ *o* has$_{t^F}\Phi$.

(PerST^{M}) Perduring object *o* has Φ at t^F in Minkowski spacetime $=_{df}$ *o*'s t^F-part has Φ.

(ExdST^{M}) Exduring object *o* has Φ at t^F in Minkowski spacetime $=_{df}$ *o*'s t^F-counterpart has Φ.

To illustrate, consider the path of a 10-meter pole in Minkowski spacetime. In the rest frame of the pole F_0 its length is 10 meters, the pole's *proper* length. In the reference frame *F*, uniformly moving in the direction of the pole, this length is Lorentz-contracted to 5 meters. This effect is a spacetime not a dynamic phenomenon and is explained by making precise what is involved in attributing length to an extended object, such as our pole, in a given perspective, or reference frame. Clearly, it involves taking the difference of the pole's ends' coordinates in that frame. These coordinates must obviously refer to the *same* time. Put another way, the events of taking the measurements of these coordinates must be simultaneous and, hence, belong to the same time hyperplane in the reference frame under consideration. Geometrically, the sought-for length is just the length of the t^F-slice through the pole's path. Not surprisingly, it turns out to be different from the proper length of the pole. Ascription of length and of many other physical properties to objects must therefore be relativized to the two-parameter index 't^F'. The endurantist, the perdurantist and the exdurantist discharge this task in their characteristic ways.

On endurantism, the pole is a 3D entity multilocated at all t^F-slices of its path and any such intersection features the full set of properties the pole has at a corresponding time in a given frame, including its length. On (Minkowskian) Relationalism, the pole comes to have the property of being 5 meters long at t^F and 10 meters long at t^{F_0} by bearing the relation *5-meter-long-at* to t^F and *10-meter-long-at* to t^{F_0}.[37] On (Minkowskian) Indexicalism, the pole accomplishes the same task

[37] Or, in a more precise analysis, to $\boldsymbol{o}_{\perp t^F}$ and $\boldsymbol{o}_{\perp t^{F_0}}$.

by exemplifying two time-indexed properties, *5-meter-long-at-*t^F and *10-meter-long-at-*t^{F_0}. On (Minkowskian) Adverbialism, the pole possesses the simple property *5-meter-long* in the t^F-ly way, and another such property, *10-meter-long*, in the t^{F_0}-ly way.

On perdurantism, the pole is a 4D entity located at its path and having a distinct momentary t^F-part at each t^F-slice through its path. Saying that the pole is 10 meters long at t^{F_0} and 5 meters long at t^F is made true by the pole's t^{F_0}-part having the former property and its t^F-part having the latter one simpliciter.

On exdurantism, the pole is a 3D entity multiLOCATED at t^F-slices through its path, in virtue of having a t^F-counterpart at each such slice. The pole is 10 meters long at t^{F_0} and 5 meters long at t^F courtesy of its t^{F_0}- and t^F-counterparts, which have these respective lengths simpliciter.

5. FLAT AND CURVED ACHRONAL REGIONS IN MINKOWSKI SPACETIME

In the generic spacetime framework introduced in Section 2, LOCATIONS of persisting objects were indexed to arbitrary achronal regions. The adaptation of the general definitions of the different modes of persistence (and of other important principles, such as Achronal Universalism) to Minkowski spacetime in Section 4 was based on the assumption[38] that persisting objects and their parts are LOCATED (and, consequently, have properties) at *flat* achronal regions representing, in special relativity, moments of time in inertial reference frames. Let us explicitly refer to this assumption as FLAT:

(FLAT) In the context of discussing persistence in Minkowski spacetime it is appropriate to restrict the LOCATIONS of persisting objects and their parts to flat achronal regions representing subsets of moments of time in inertial reference frames.

Initially one might be inclined to reject FLAT on rather general metaphysical grounds. Consider a *non*-flat achronal slice $\boldsymbol{o}_\perp$ of object *o*'s path in Minkowski spacetime. How could *o* (if *o* endures or exdures), or one of *o*'s diachronic parts (if *o* perdures), *fail* to be LOCATED at $\boldsymbol{o}_\perp$? In other words, how could $\boldsymbol{o}_\perp$ fail to "contain" *o* (or one of *o*'s diachronic parts)? After all, $\boldsymbol{o}_\perp$ is an achronal slice of *o*'s path and is matter-filled; therefore it must contain *something*! And what could this "something" be except *o* or one of *o*'s diachronic parts?

This general line of thought should be resisted (cf. Gilmore, 2006: 210–211), because in turns on conflating the notion of an achronal region's being a LOCATION of *o* (or one of its achronal parts) with the notion of an achronal region's being "filled with achronal material components of *o*." A region may satisfy the latter property without satisfying the former. Imagine Unicolor, a persisting object one of whose essential properties is to be *uniformly colored*. Suppose further that Unicolor uniformly changes its color with time in a certain inertial reference frame *F*.

[38] Shared by a number of other writers; see, in particular, Sider (2001: 59, 84–86), Rea (1998), Sattig (2006: Sections 1.6 and 5.4). Gilmore, who held this assumption in his earlier work (Gilmore, 2004), appears to have abandoned it later (see, in particular, Gilmore, 2006). Unlike Gibson and Pooley (2006), however, he does not offer any specific criticism of the assumption.

Consider an achronal slice of Unicolor's path, flat or not, that crisscrosses hyperplanes of simultaneity in F. Whatever (if anything) is LOCATED at such a slice is not uniformly colored and, hence, must be distinct from Unicolor, even though it is filled with the (differently colored) achronal material components of Unicolor.

This shows that general metaphysical considerations are not sufficient to reject FLAT. But notice that the property of being *uniformly colored* used in the above example is itself grounded in a prior concept of spatial or achronal *uniformity*, which, in turn, presupposes that flat achronal regions of Minkowski ST are somehow physically privileged in the context of SR. In a recent work Gibson and Pooley (2006: 160–165) have argued that they are not, thereby presenting a more pointed objection to FLAT. Their objection also raises important methodological questions about the relationship between physics and metaphysics. Below I consider and respond to Gibson and Pooley's objection and, in the course of doing it, address the methodological concerns brought to light in their critique of FLAT.

In Gibson and Pooley's view, the tendency to "frame-relativize" in the manner of FLAT and other similar assumptions, which is adopted unreflectively by several authors discussing persistence in the context of Minkowski spacetime (see note 38), represents a relic of the classical worldview and stands in the way of taking relativity seriously. While inertial frames of reference (i.e., spacetime coordinate systems adapted to them) are geometrically privileged and, therefore, especially convenient for describing spatiotemporal relations in Minkowski spacetime, this does not give them any distinguished metaphysical status. Accordingly (and contrary to FLAT), no such status should be granted to flat achronal regions in Minkowski spacetime. Thus Gibson and Pooley:

> From the physicist's perspective, the content of spacetime is as it is. One can choose to describe this content from the perspective of a particular inertial frame of reference (i.e., to describe it relative to some standard of rest and some standard of distant simultaneity that are optimally adapted to the geometry of spacetime but are otherwise arbitrary). But one can equally choose to describe the content of spacetime with respect to some frame that is not so optimally adapted to the geometric structure of spacetime, or indeed, choose to describe it in some entirely frame-independent manner (Gibson and Pooley, 2006: 162).
>
> ...
>
> More significantly, one surely wants a definition [of a notion relevant to characterizing a particular mode of persistence in spacetime—Y.B.] applicable in the context of our best theory of space and time, general relativity. While this theory allows spacetimes containing flat spacelike regions, generic matter-filled worldtubes will have *no* flat maximal spacelike subregions. The obvious emendation, therefore, is simply to drop clause (iv) [i.e., FLAT or some analogous assumption—Y.B.] (Ibid., 2006: 163).

These remarks contain two distinct points, and both raise important questions. The first point—that inertial reference frames and flat regions in Minkowski spacetime are privileged only geometrically and not physically and, therefore, do not warrant ascribing to them any metaphysical significance in the context of ques-

tions about persistence—appears to derive its force from a crucial lesson of the contemporary methodology of spacetime theories: that the choice of a local coordinate system is completely arbitrary and has no bearing whatsoever on the content of a particular spacetime theory.[39] Any such theory—Newtonian mechanics, classical electrodynamics or special relativity—can be formulated in any coordinate system. Moreover, such a formulation can always be made covariant with respect to arbitrary local coordinate transformations, at the cost of making it less elegant. For example, Newtonian mechanics of free particles in Galilean spacetime can be stated in terms of a set of geometrical objects on the manifold:[40] an affine connection D, a covariant vector field dt, and a two-rank symmetric tensor h, satisfying the following field equations:

$$R^{\mu}{}_{\nu\lambda\kappa} = 0, \quad t_{\mu;\nu} = 0, \quad h^{\mu\nu}{}_{;\lambda} = 0, \quad h^{\mu\nu}t_{\mu} = 0$$

and the equations of motion:

$$\frac{d^2x_{\mu}}{du^2} + \Gamma^{\mu}_{\lambda\kappa}\frac{dx_{\lambda}}{du}\frac{dx_{\kappa}}{du} = 0,$$

where u is a real-valued parameter and ';' denotes covariant differentiation. The above represents the statement of the theory in arbitrary local coordinate systems. As Gibson and Pooley note, a spacetime theory such as Newtonian mechanics can also be given a coordinate-free formulation:

$$K = 0, \quad \bar{D}(dt) = 0, \quad \bar{D}(h) = 0, \quad h(dt, w) = 0$$

where w is a covariant real vector field in the cotangent space defined at a given spacetime point.

It turns out that there is a special sub-class of *inertial* coordinate systems—defined locally by $\Gamma^{\mu}_{\lambda\kappa} = 0, t_{\mu} = (1,0,0,0)$, and $h^{\mu\nu} = \delta^{\mu\nu}$ for all μ and ν except $\mu = \nu = 0$, while $h^{00} = 0$—in which the equation of motion takes the familiar form of Newton's First Law:

$$\frac{d^2x_{\mu}}{dt^2} = 0.$$

Although this fact obviously has enormous practical significance: it allows us to use a simple expression of Newton's First Law in a great variety of practical applications, the fact that such frames exist has no physical importance. Indeed, suppose a certain particle performs a non-inertial motion. One could then associate with it a series of instantaneous rigid Euclidean systems, for which $\Gamma^{\mu}_{\lambda\kappa}$ will not vanish, and recover the equation of motion (Friedman, 1983: 83):

$$\frac{d^2x_{\mu}}{dt^2} + a^{\mu} + 2\Omega^{\mu}_{\nu}\frac{dx_{\nu}}{dt} = 0,$$

where $x_{\mu}(t) \equiv x_{\mu} \circ \sigma(t)$ is a family of continuous and differentiable real functions of the time-parameter t,

$$a^{\mu} \equiv \frac{d^2x_{\mu}}{dt^2} + \Gamma^{\mu}_{\lambda\kappa}\frac{dx_{\lambda}}{dt}\frac{dx_{\kappa}}{dt}$$

[39] See, for example, Friedman (1983: Section II.2).

[40] My outline of this example follows Friedman (1983: 87–94).

is the acceleration and $\Omega^{\mu}_{\nu} \equiv \Gamma^{\mu}_{0\nu} = \Gamma^{\mu}_{\nu 0}$ is an antisymmetric rotation matrix. This equation of motion features the inertial force a^{μ} associated with the acceleration of the rest frame of the particle and the Coriolis force $2\Omega^{\mu}_{\nu}\frac{dx_{\nu}}{dt}$ associated with its rotation.

The point to note here is that the presence of straight non-achronal "position lines," which allow one to identify spatial positions at different times in perspectives associated with inertial coordinate systems, has no physical consequence. Based on this point, one could argue that position in space, as defined in a *given* inertial frame, is a rather thin notion that hardly bears the weight attributed to it in many metaphysical discussions—even in the context of classical physics.

And things get worse. Even in *that* context, one can choose to "geometrize away" gravitational forces by incorporating the gravitational potential into the affine connection (Friedman, 1983: 100):

$$\Gamma^{\mu'}_{\lambda\kappa} = \Gamma^{\mu}_{\lambda\kappa} + h^{\mu\lambda}\Phi_{;\lambda}t_{\lambda}t_{\kappa}$$

at the cost of making the classical spacetime non-flat (i.e., by making it curved).[41]

This example shows that, *even in the classical context*, the presence of a well-defined family of straight diachronic position lines and the usual assumption that the spacetime as a whole is flat have no physical significance. Does this mean that one should ban familiar notions, such as *same place over time in a given inertial frame*, from philosophical discussions tailored to the classical context, simply because inertial frames and straight achronal lines enjoy no special status at the fundamental level of physical description?

Hardly so. Banning such notions would deprive one of many useful resources in the situation where such resources are *available*. Note that the issue does not concern the retention of the notion of *sameness of place over time*, period (even the classically-minded metaphysician can be convinced that the latter notion *is* meaningless), but only the significance of the notion of sameness of place over time *in an inertial frame*. This notion provides resources for imposing on spacetime a global coordinatization and assigning to such coordinatization various conceptual roles. It would appear that the metaphysician should feel free to make use of the familiar concept of sameness of place across time (against the backdrop of a particular inertial frame)—as long as such a concept is definable—even if physics, in the end, denies distinction to inertial frames.

Two facts seem to be relevant here: (i) that global inertial coordinate systems are *available* (despite the lack of physical importance) and (ii) that their availability allows one to *minimize revision* of the existing ontological vocabulary. The above brief excursus into Newtonian mechanics should serve to support (i). (ii), on the other hand, raises more general considerations having to do with philosophical methodology.

It is a well-known fact that most contemporary discussions in fundamental ontology[42] continue to be rooted in the "manifest image of the world" and ignore

[41] We shall not pursue this further. See Friedman (1983: 95–104) for details.

[42] That is, discussions of such issues as time, persistence, material composition, the nature of fundamental properties and laws, etc.

important physical developments, which have rendered many common-sense notions untenable and obsolete. Attempts to bring physical considerations to bear on issues in fundamental ontology, such as those discussed in this chapter, are still very rare. This persistent self-isolation of contemporary metaphysics from science may prompt at least two different reactions from philosophers who are wary of "armchair philosophical speculation." One may be tempted to reject such speculation, root and branch, and adopt the following attitude: let physics tell us what the world is like and then let the "metaphysical chips" fall where they may. It is unclear whether any part of the contemporary metaphysical agenda would survive such a treatment. But it is equally unclear whether any consistent world view could emerge from it. Science is an open-ended enterprise which is becoming increasingly fragmented. The same is true of any particular scientific discipline, such as physics. The question of what parts of contemporary fundamental physics could contribute safe and reliable components to the foundations of an overall world view is a highly complex question, which may not have a good answer.

This suggests a different attitude. One may admit that the prolonged mutual alienation of metaphysics and physics is unfortunate but insist that both have some value in their *current* state, and could therefore benefit from gradual rapprochement. It should be clear that the present chapter follows the second course. It should also be clear that this course brings with it certain limitations. One of them has to do with the choice of the physical theory (or theories) under consideration. Given the open-ended nature of physics any physical theory is likely to be false. But one hopes that some theories are good approximations to the truth, and to the extent that they are, adapting existing metaphysical views to them is valuable. The scope of the present consideration does not go beyond special relativity. This represents a particular choice and brings with it quite obvious restrictions.

Even more important, when engaged in extending an existing metaphysical debate to a new physical framework one confronts non-trivial judgment calls at many turns, when it becomes clear that some familiar notions must be abandoned, others modified, while others can be kept more or less intact. Usually there is more than one way to "save the philosophical appearances," but the decision as to what "intuitions" must be retained at the expense of others is difficult because one is now swimming in uncharted waters. In the end, it is the entire resulting systems and their performance across a variety of theoretical tasks that must be compared. I submit that the only reasonable regulative maxim to be imposed on physically-informed metaphysical theorizing should be stated in terms of *Minimizing the Overall Ontological Revision* (MOOR). Vague as it is, its role could be favorably compared to Quine's famous criteria of "conservatism," "the quest for simplicity" and "considerations of equilibrium" affecting the "web of belief as a whole":

(MOOR) In adapting a metaphysical doctrine to a physical theory one should seek to minimize the degree of the overall ontological revision.

As we depart from the "comfort zone" of the classical world view, the degree and extent of the "overall ontological revision" become progressively up for grabs, which makes MOOR increasingly wholesale and non-specific. But as indicated above, any alternative to MOOR would amount to rejecting the entire agenda of contemporary metaphysics. I should emphasize that the latter is not what Gibson and Pooley undertake to do in the above-quoted work (Gibson and Pooley, 2006). Having noted that they have "a lot of sympathy" for the view that "the project of reconstructing relativistic version of familiar non-relativistic doctrines [may be] horribly misguided,"[43] they "nonetheless think that it is worthwhile to engage with attempts to square the familiar doctrines with relativity" (ibid., 2006: 157–8). Such attempts, I recommend, must be guided by something like MOOR.

Returning (finally) to FLAT, I contend that it conforms to the spirit of MOOR quite well. Indeed FLAT employs structures (viz., global flat hypersurfaces) that are (i) available in Minkowski spacetime, (ii) widely used in physics, and (iii) are indispensable to extending the important notions of moment of time and momentary location of an object or its part (in a given reference frame) to the special relativistic framework. In this respect, FLAT is on a par with the license to attribute metaphysical importance to a family of straight positions lines in classical spacetime despite the fact that, at bottom, straight diachronic lines do not enjoy (*even* in Galilean spacetime) any physical privilege over curved diachronic lines. The important facts are that (i) straight lines are *definable* in that context and that (ii) without their presence, the notion of "place over time in a given frame" would get completely out of touch with any familiar notions. For similar reasons, global hyperplanes can be assigned important metaphysical roles in Minkowski spacetime. First, they are easily definable as such. Second, if they lose their privilege over arbitrary achronal hypersurfaces vis-à-vis issues of persistence, the notion of *momentary* location of a persisting object—and, with it, the host of other notions tied up to momentary location, such as momentary shape, momentary achronal composition, and the like—would lose much of their ground and would be hard to connect to any familiar concepts. They would become too remote to perform any meaningful function in a metaphysical debate.

I conclude that FLAT is justified in the context of Minkowski spacetime. But I fully agree with Gibson and Pooley that it is not appropriate for *general* relativistic spacetime, where matter-filled flat achronal regions are *not available*. Since that context has no place for global "moments" of time and "momentary" locations, the connection with the familiar set of notions is severed anyway and there is no pressure to align other concepts with them. In general relativistic spacetime it is only natural to regard any achronal slice of an object's path as a good candidate for the object's (or its part's) location—if one thinks that the notion of location continues to make any sense there.[44]

[43] "Should we not start with the relativistic world picture and ask, in that setting and without reference to non-relativistic notions, how things persist?" (Gibson and Pooley, 2006: 157–8).

[44] My consideration is restricted, for the most part, to Minkowski spacetime of special relativity, which, for the purpose of discussion, is taken to be a good approximation of the spacetime of our real world. Even so, the issue of the status of curved hypersurfaces in Minkowski spacetime is more interesting than it might appear. Some facts about such hypersurfaces are non-trivial and notable in their own right. For discussion of one such fact, see Balashov (2005: Section 9 and Appendix.).

ACKNOWLEDGEMENTS

My greatest debt is to Maureen Donnelly and Cody Gilmore for spotting multiple errors in several consecutive drafts and for their very helpful suggestions. The remaining defects are solely my responsibility. Thanks are due to an anonymous referee for insightful critique of an earlier draft. Work on this chapter was supported by a Senior Faculty Grant from the University of Georgia Research Foundation.

REFERENCES

Balashov, Y., 1999. Relativistic objects. Noûs 33, 644–662.
Balashov, Y., 2000a. Enduring and perduring objects in Minkowski space-time. Philosophical Studies 99, 129–166.
Balashov, Y., 2000b. Persistence and space-time: Philosophical lessons of the Pole and Barn. The Monist 83, 321–340.
Balashov, Y., 2005. Special relativity, coexistence and temporal parts: A reply to Gilmore. Philosophical Studies 124, 1–40.
Balashov, Y., 2007. Defining 'Exdurance'. Philosophical Studies 133, 143–149.
Bittner, Th., Donnelly, M., 2004. The mereology of stages and persistent entities. In: Lopez de Mantaras, R., Saitta, L. (Eds.), Proceedings of the European Conference of Artificial Intelligence. IOS Press, pp. 283–287.
Butterfield, J., 2006. The rotating discs argument defeated. The British Journal for the Philosophy of Science 57, 1–45.
Casati, R., Varzi, A.C., 1999. Parts and Places: The Structures of Spatial Representation. The MIT Press, Cambridge, MA.
Crisp, Th., Smith, D., 2005. 'Wholly Present' defined. Philosophy and Phenomenological Research 71, 318–344.
Friedman, M., 1983. Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science. Princeton University Press, Princeton.
Gibson, I., Pooley, O., 2006. Relativistic persistence. In: Hawthorne, J. (Ed.), Metaphysics. In: Philosophical Perspectives, vol. 20. Blackwell, Oxford, pp. 157–198.
Gilmore, C., 2004. Material Objects: Metaphysical Issues. Princeton University Dissertation.
Gilmore, C., 2006. Where in the relativistic world are we? In: Hawthorne, J. (Ed.), Metaphysics. In: Philosophical Perspectives, vol. 20. Blackwell, Oxford, pp. 199–236.
Gilmore, C., 2007. Time travel, coinciding objects, and persistence. In: Zimmerman, D. (Ed.), Oxford Studies in Metaphysics, vol. 3. Clarendon Press, Oxford, pp. 177–198.
Haslanger, S., 2003. Persistence through time. In: Loux, M.J., Zimmerman, D. (Eds.), The Oxford Handbook of Metaphysics. Oxford University Press, Oxford, pp. 315–354.
Hawley, K., 2001. How Things Persist. Clarendon Press, Oxford.
Heller, M., 1990. The Ontology of Physical Objects: Four-dimensional Hunks of Matter. Cambridge University Press, Cambridge.
Hudson, H., 2001. A Materialist Metaphysics of the Human Person. Cornell University Press, Ithaca.
Hudson, H., 2006. The Metaphysics of Hyperspace. Oxford University Press, Oxford.
Lewis, D., 1986. On the Plurality of Worlds. Blackwell, Oxford.
MacBride, F., 2001. Four new ways to change your shape. The Australasian Journal of Philosophy 79, 81–89.
McKinnon, N., 2002. The endurance/perdurance distinction. The Australasian Journal of Philosophy 80, 288–306.
Merricks, T., 1999. Persistence, parts and presentism. Noûs 33, 421–438.
Parsons, J., 2000. Most a four-dimensionalist believe in temporal parts? The Monist 83, 399–418.

Parsons, J., 2007. Theories of location. In: Zimmerman, D. (Ed.), Oxford Studies in Metaphysics, vol. 3. Clarendon Press, Oxford, pp. 201–232.
Rea, M., 1998. Temporal parts unmotivated. Philosophical Review 107, 225–260.
Simons, P., 1987. Parts. A Study in Ontology. Clarendon Press, Oxford.
Sattig, T., 2006. The Language and Reality of Time. Clarendon Press, Oxford.
Sider, T., 2001. Four-Dimensionalism. An Ontology of Persistence and Time. Clarendon Press, Oxford.
Sklar, L., 1974. Space, Time, and Spacetime. University of California Press, Berkeley.
Varzi, A., 2003. Naming the stages. Dialectica 57, 387–412.
Zimmerman, D., 1996. Persistence and presentism. Philosophical Papers 25, 115–126.

CHAPTER 5

Is Spacetime a Gravitational Field?

Dennis Lehmkuhl*

Abstract

I point out that the often voiced claim that in the general theory of relativity (GR) geometry and gravity are 'associated' with each other can be understood in three different ways. The *geometric interpretation* asserts that gravity can be reduced to spacetime geometry, the *field interpretation* claims that the geometry of spacetime can be reduced to the behaviour of gravitational fields, and the *egalitarian interpretation* affirms that gravity and spacetime geometry are conceptually identified. I investigate different versions of each interpretation and argue that an egalitarian interpretation is the one most faithful to the formalism of GR. I then briefly review two rival theories of GR, Brans–Dicke theory and Rosen's first bimetric theory, thereby showing that this is not the case for every modern theory of gravity, and that hence the one-to-one correspondence between geometry and gravity is a peculiar feature of GR.

1. INTRODUCTION

Particle physicists tell us that there are five forces: the electric force, the magnetic force, the weak force, the strong force—and the gravitational force. Three of these forces, we are rightly told, have already been unified: the electric and magnetic forces by Faraday, Ampère and Maxwell; the electromagnetic and weak forces by Glashow, Salam and Weinberg. Furthermore, the unification of the electroweak and strong forces is supposed to be underway, giving us a "Grand Unified Theory (GUT)". The final goal is then supposed to be a "theory of everything": a unification of the electro-magnetic-weak-strong forces (which govern the universe on

* Oriel College, Oxford University, UK
E-mail: dennis.lehmkuhl@philosophy.ox.ac.uk

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00005-3
© Elsevier BV
All rights reserved

small scales) with the gravitational forces (which govern the universe on large scales).[1] But is gravity a force like the others? *Is it a force at all?*[2]

It is often claimed that the core of the theory of general relativity (GR) is that gravitational phenomena, that would otherwise be explained by the action of forces, are just a manifestation of spacetime geometry. Indeed, Einstein's idea of *associating* gravity with the geometry of spacetime was surely one of the most beautiful ideas in the history of physics. However, it remains controversial *in what sense* gravity and geometry are associated with each other.

There are two obvious possibilities about what it could mean to say that GR associates gravity with the geometry of spacetime. Both of them are present in the literature, but it is often not clear in how far one is superior to the other. In a nutshell, one of them sees gravity as a 'manifestation' of spacetime geometry, while the other sees spacetime geometry as a manifestation of gravitational fields. I will call the former the *geometric interpretation* of GR, or the assertion that GR is the 'geometrization of gravity', whereas the latter will be called the *field interpretation*, or the claim that GR is the 'gravitization of geometry'. Both of them seem to aim for a *reduction* of either gravity or the geometry of spacetime, respectively, to the other.

There is a third possible position, which I will call the *egalitarian interpretation* or *egalitarianism* for short. Variants surface in the texts of proponents of both geometric and field interpretation. A weak version of egalitarianism claims that the formalism of GR has both geometric and field significance, a strong version goes further and claims that geometry and gravity are conceptually identified within the theory, making them two names for one and the same 'thing'. My aim in this text is to discuss the viability of different versions of the egalitarian interpretation of GR—and to find out which constitute a genuine alternative position that may have merits of its own.

I will start out in Section 2 by giving a first characterization of the three positions of how gravity and geometry relate to each other. One might object that we need to know what precisely we mean by 'geometry' and 'gravity' in order to start investigating how they relate to each other. I do not agree: discussing how the two concepts relate to each other according to the formalism of GR is surely a step *towards* understanding what these concepts mean (in GR), rather than something that has to be presupposed. In Section 3, I will define what it means for a mathematical object to have 'geometrical significance' and/or 'gravitational significance', although I will not claim that these definitions are sufficient to give a definition of the more fundamental terms 'geometry' and 'gravity'. In Section 4, I will review the discussion of which mathematical object in GR represents physical gravitational fields, and use it in order to clarify which parts of the formalism have gravitational and/or geometrical significance. With the results of this discussion in hand, I will then return in Section 5 to the initial characterization of

[1] See Maudlin (1996) for an analysis of the unificationary pursuit in modern physics, comparing the different kinds of unification in electromagnetism, general relativity and gauge theories.

[2] Many physicists would speak of electromagnetic, weak and strong *interactions*, and use the term 'force' only in a metaphoric way in order to speak about these interactions. The question is then: is the gravitational interaction an interaction that is not just different from the others in the sense that weak and strong interaction are different from each other, but a different *kind* of interaction?

geometric, field and egalitarian interpretation, showing how the results obtained so far point to different variants of the three positions. In particular, I will point out that there are three versions of egalitarianism—a weak, a moderate and a strong version. Then, in Section 6, I will discuss two rival theories of GR, Brans–Dicke theory and Rosen's first bimetric theory, in order to find out whether the possibility of an egalitarian interpretation is peculiar to GR, or rather a generic feature of modern theories of gravity.

2. ASSOCIATING GEOMETRY AND GRAVITY

2.1 Associating as reducing: the geometric and the field interpretation

As far as I know, the distinction between the two positions that I call the geometric and the field interpretation[3] was first pointed out by Hans Reichenbach—but his distinction was not really taken up. Reichenbach writes:[4]

> [The] universal effect of gravitation on all kinds of measurement instruments defines therefore a single geometry. In this respect we may say that gravitation is *geometrized*. We do not speak of a change produced by the gravitational field in the measuring instruments, but regard the measuring instruments as "free from deforming forces" in spite of the gravitational effects.
>
> However, we have seen that for geometry, as for all other phenomena, we must pose the question of causation. [...] In this sense we must ascribe to the gravitational field the reality of a force field. We regard this force field as the cause of the geometry itself, not as the cause of the disturbance of geometrical relations. [...] *[I]t is not the theory of gravitation that becomes geometry, but it is geometry that becomes an expression of the gravitational field*.

One may criticise Reichenbach's remarks on various grounds, in particular with regard to his notion of causation. Indeed, I think that neither his way of putting forward a geometric interpretation of the formalism, nor his way of formulating a field interpretation, are the most sensible ones, as will become evident below.[5]

But for now, the important point is that Reichenbach was the first to make a distinction between viewing either geometry or gravity as more fundamental, and the idea of one being a manifestation or consequence of the other. Apart from being realist positions, both the geometric interpretation and the field interpretation have in common that they seem to be *reductionist* positions.

Carlo Rovelli likewise distinguishes between two interpretations of GR. He writes:[6]

[3] Note that both are realist positions: they agree that both 'geometry' and 'gravity' refer to something in 'the world', they just differ about the way the terms do so.

[4] See Reichenbach (1957, p. 256). That they are actually two distinct and—for Reichenbach—mutually exclusive options can be seen most clearly from the very last sentence of the quotation below (his emphasis).

[5] For example, Reichenbach presupposes that the gravitational field has to be a *force* field in order not to be geometrized. I will later argue that this is not true: one can regard the gravitational field as a physical field that is *not* reduced to geometry even though it is not seen as a force field.

[6] See Rovelli (1997, pp. 193–194). Rovelli claims that the decision between these two possibilities is just "a matter of taste, at least as long as we remain within the realm of nonquantistic and nonthermal general relativity." In Section 5, I will

Einstein's identification between gravitational field and geometry can be read in two alternative ways:

i. as the discovery that the gravitational field is nothing but a local distortion of spacetime geometry; or
ii. as the discovery that *spacetime geometry is nothing but a manifestation of a particular physical field*, the gravitational field.

Arguably, the geometric and the field interpretation can be cashed out in different ways, making them families of positions rather than single positions. However, all variants will have in common that they take either geometry or gravity to be more fundamental than the other.

On the geometric side, Misner, Thorne and Wheeler speak of gravitation being a *manifestation* of the curvature of spacetime,[7] Hartle speaks of gravitational phenomena *arising* from spacetime curvature,[8] whereas Wald affirms that gravity is an *aspect* of spacetime structure.[9]

Proponents of the field interpretation have a similar vocabulary: Reichenbach speaks of geometry being an *expression* of the gravitational field,[10] Bergmann speaks of the gravitational field *forming* the structure of spacetime,[11] and Einstein himself even speaks of Minkowski spacetime as being a *special type* of gravitational field.[12]

Of course, it is still widely held that Einstein himself regarded GR as a geometrization of gravity. That is surely not adequate, for in a letter to Barret he explicitly says that he does "not agree with the idea that the general theory of relativity is geometrizing Physics or the gravitational field", adding remarks that clearly show that he regards field theoretical concepts as more fundamental than geometrical concepts.[13] Similar remarks can be found in his autobiographical notes.[14]

Both the geometric and the field interpretation of GR can indeed be seen as interpretations in the classical sense: they do not require an immediate change of the formalism itself.[15] However, taking either of the two stances does strongly influence the picture of GR one has, to an extent that may cause different sets of questions to seem important or even to be seen as answerable. For example, a

argue that this constitutes a moderate form of egalitarianism. Adopting a strong form of egalitarianism makes it more than a matter of taste.

7 See Misner et al. (1973, p. 304).

8 See Hartle (2003, p. 107).

9 See Wald (1984, p. 67).

10 See Reichenbach (1957, p. 256), i.e., the quote cited above, where he not only lists the two possibilities but also endorses his version of the field interpretation.

11 See Bergmann (1976, p. 245).

12 See Einstein (1917; English edition: p. 155). Although the book was written in 1917, the appendix which contains the quote referred to was written in 1954. Note that Einstein speaks of Minkowski spacetime as it occurs in GR, namely as a solution of the Einstein field equations—the Minkowski spacetime of special relativity is of course not connected to gravitation.

13 Einstein in a letter to Barret from 1948, as cited in Stachel's preface to the proceedings of the Andover conference on the "Foundations of Space-Time Theories" in 1977 (Earman et al., 1977, p. ix).

14 See Einstein (1949, p. 56/57).

15 A counterexample to an interpretation in the classical sense is the reformulation of GR in terms of spin-2-fields (gravitons) on flat spacetime—which I will nevertheless briefly discuss in Section 5, given that it seems a very natural route to take if one starts out with a 'pure' interpretation of GR in terms of gravitational fields rather than of spacetime geometry. In quantum mechanics, the Bohmian account, as well as the GRW account, are in this sense not 'pure' interpretations.

commitment to the field interpretation makes it very natural to see GR in strong analogy to Maxwell's theory of electromagnetism, and hence to first search for solutions of the Einstein equation which correspond to gravitational waves. (This is exactly what Einstein did in (1916), only one year after having found the Einstein equation.[16]) On the other hand, a commitment—or even just a sympathy—to the geometric interpretation makes it natural to search for solutions describing spacetime as a whole. (Such solutions are, e.g., the Friedmann–Lemaitre–Robertson–Walker solutions to the Einstein equation.) One might argue that in delivering different heuristics, both interpretations (and probably interpretations in general) are capable of delivering 'second order changes' to the formalism.[17]

In any case, we have two positions, one of which regards gravity as being reducible to spacetime geometry, while the other regards spacetime geometry as being reducible to the behaviour of gravitational fields. These are just initial and rather vague characterizations: I will come back to different precise variants of both interpretations in Section 5. But first there is a third possible position: the egalitarian interpretation.

2.2 Associating as identifying: the egalitarian interpretation

The results found by starting out from both geometric and field interpretation will be valuable. Hence, if one of the main motivations for a *philosophical* interpretation of GR is to point to future *physical* results, then one might ask: why should we restrict ourselves to the potential fruitfulness of only one interpretation?

One might answer: because physical fruitfulness is not the only reason why we seek a philosophical interpretation of GR. The other reason is that we want to know what GR tells us about *nature*, and it simply cannot be that geometry is both an aspect of gravity and gravity an aspect of geometry. 'A is an aspect of B', or 'A is reducible to B', is evidently not symmetric.

Or so it might seem. Unless we can take what almost looks like a cheap solution: it is not the case that one is conceptually *reduced* to the other; the two are rather conceptually *identified*. I will call this view the *strong egalitarian interpretation*, or (for the time being) *egalitarianism* for short.

So far this possibility is just a pious hope: we would like to have our cake and eat it too. For if gravitational field and spacetime geometry were after all essentially one and the same 'thing', it would be perfectly consistent to switch back and forth between the heuristics of the field interpretation and those of the geometric interpretation: after all, gravity and geometry would be equally fundamental through being identified with each other. ('A is reducible to B' is trivially symmetric given that $A = B$.)

Indeed, we find statements of an egalitarian flavor in many texts, often even in those of passionate proponents of one of the two standard interpretations. Mis-

[16] The paper contains a substantial calculational error, which Einstein corrected one and a half years later in Einstein (1918). For a very clear exposition of how central it was for the development of GR for Einstein to see the theory as being mainly about gravitational fields rather than about the geometry of spacetime see Renn and Sauer (2006).

[17] Belot (1998) makes points kindred in spirit. Also, in Belot (1996), he gives general arguments for the need for an interpretation of classical GR, in particular in facing the task of constructing a theory of quantum gravity.

ner, Thorne and Wheeler, allegedly archbishops of the geometric interpretation, write:[18]

> [N]owhere has a precise definition of the term "gravitational field" been given—nor will one be given. Many different mathematical entities are associated with gravitation: the metric, the Riemann curvature tensor, the Ricci curvature tensor, the curvature scalar, the covariant derivative, the connection coefficients etc. Each of these plays an important role in gravitation theory, and none is so much more central than the others that it deserves the name "gravitational field." Thus it is that throughout this book that the terms "gravitational field" and "gravity" refer in a vague, collective sort of way to all of these entities. Another, equivalent term used for them is the "geometry of spacetime."

Anderson also claims that "geometry and gravitation [are] one and the same thing",[19] but he does not elaborate. Even Feynman, supposedly the nemesis of the geometric interpretation, writes that he wants "to understand how gravity can be both geometry and field".[20]

I will later (in Section 5) argue that there are different forms of egalitarianism, some of which are indeed compatible with both geometric and field interpretations. This also explains why we find egalitarian remarks in the texts of proponents of both standard interpretations. However, strong egalitarianism will remain an alternative interpretation—and indeed a very promising one, as we will see. But first, I will have to make precise what it means for a given mathematical object to have geometrical and/or gravitational significance (Section 3), and clarify what should be reasonably understood under the term 'gravitational field' in the context of standard GR (Section 4); hence arguing that it is not necessary to use the term in a "vague, collective sort of way".

3. GEOMETRICAL AND GRAVITATIONAL SIGNIFICANCE

Which of the three interpretations is the one most faithful to the formalism? Before I can start to answer this question, two definitions are in order.

A mathematical object has *geometrical significance* if it gives us an account of geometrical phenomena, and hence is needed in a theory to represent or model aspects of physical spacetime geometry. Note that this notion of a mathematical object having *geometric* significance is a weaker condition than Brown's (2005) notion of a mathematical object having *chronogeometric* or chronometric significance. According to Brown, a mathematical entity has chronometric significance if it can be related to measurements we can perform with physical rods and clocks, even if only under idealized conditions. I will take chronometric significance as a sufficient, but not necessary condition for geometric significance in the sense defined above.

[18] See Misner et al. (1973, p. 399).
[19] See Anderson (1967, p. 334).
[20] See Feynman (1995, p. 113), and Section 5 for a full quote in the context of the spin-2 approach to GR.

The reason is not that it is controversial whether clocks and rods should play a fundamental role in GR;[21] chronometricity in Brown's sense only demands a *relation* to rods and clocks to be possible in order for a given mathematical object to have geometric significance, not that rods and clocks themselves are fundamental in some sense. But if we model physical spacetime, for example, by a purely affine mathematical geometry, i.e. a geometry in which the metric g_{ab} plays no fundamental role, but only the connection $\Gamma^a{}_{bc}$ (cf. below), we will still want to attribute geometrical significance to the connection, as long as it gives us an account of (some) geometrical phenomena. Paradigm examples for geometrical phenomena do of course include phenomena involving the behaviour of rods and clocks (time dilation, length contraction, simple distance measurements), but also e.g. the geodesic motion of test bodies, which are describable solely in terms of the connection.

Likewise, a mathematical object has *gravitational significance* if it plays a role in describing and explaining the phenomena we count as gravitational. Examples of gravitational phenomena are facts like the attraction of the earth by the sun, the fact that things fall towards earth, and that they do not fall parallel to each other. These phenomena were already regarded as being gravitational in nature before the dawn of GR—now the list has to be supplemented by phenomena like light being deflected by the sun and gravitational redshift.[22] This already shows that we may not have a complete list of gravitational phenomena; any new theory of gravitation may predict new ones, and may even be demanded to do so. But the point is that a mathematical object in a given theory has gravitational significance if it plays a role in accounting for gravitational phenomena, both old and new.[23]

Most theories of gravitation will presume mathematical objects having gravitational significance to 'couple' to the mathematical objects representing matter in a manner such that the coupling does not depend on the constitution or kind of matter (including non-gravitational fields): every kind of matter (if present) is affected by fields with gravitational significance in a universal way. However, this criterion presupposes the equivalence of gravitational and inertial mass (the weak equivalence principle, or WEP for short). The latter is today widely accepted as an empirical fact, but we surely would not want to say that presupposing it is a necessary requirement to even call a given theory a theory of gravitation.[24] Or to put it differently: even if we would discover a new sort of particle that is not in-

[21] Synge (1956, 1960) has argued that we should use *only* clocks as basic in GR. He showed that it is possible to start out with the time-like worldlines of a collection of clocks, in order to then *determine* indirectly the values of the line element ds for event pairs whose separation is space-like (Synge, 1956, p. 23). Of course, it remains true that these values of the line-element are the ones we measure with rods, but the importance of this fact is questioned by Synge's method. Ehlers, Pirani and Schild go even further. They criticize Synge in detail, finally rejecting clocks as basic tools for giving an account of spacetime geometry and proposing to use light rays and freely falling particles instead (Ehlers et al., 1972, pp. 64–65). For a comparison of the three positions (i: rods *and* clocks fundamental, ii: *only* clocks fundamental, iii: *only* light rays and freely falling particles fundamental) compare also Grünbaum (1973).

[22] See Feynman (1995, pp. 3–10) for an investigation of "the characteristics of gravitational phenomena".

[23] Whether a new phenomenon is gravitational in nature is a question to be argued for on a case by case basis, and surely not a theory-independent question.

[24] At the very least, this would be ahistorical. For there *were* theories of gravitation that did violate the WEP, and which seemed viable candidates before the equivalence of gravitational and inertial mass became widely accepted through the Eötvös experiments (Eötvös, 1889; Eötvös et al., 1922). An example of such a theory is Nordström's first theory from 1912; see Norton (1992) for an extensive review of all of Nordström's theories, with a particular emphasis on the exchange between Nordström and Einstein and the role of the equivalence of inertial and gravitational mass in the different theories.

fluenced by gravitational fields, we would not say that gravity has ceased to exist altogether.

The mathematical geometries used in general relativity and most classical unified field theories have in common that they rest on two fundamental mathematical objects, the metric tensor g_{ab} and the connection $\Gamma^a{}_{bc}$.[25] In general, the metric and connection are two independent mathematical objects, whereas a connection always defines one or more curvature tensors $R_{abc}{}^d$.[26]

The mathematical basis of GR is pseudo-Riemannian geometry. Its specifications are that

1. the metric is *symmetric*: $g_{ab} = g_{ba}$;
2. the metric has a *Lorentzian* signature, which in four dimensions amounts to sign $= (-+++)$ or sign $= (+---)$;[27]
3. the metric is *nondegenerate*: $\forall v \in V_p$: $g(v, v_1) = 0$ iff $v_1 = 0$ where V_p is the space of tangent vectors at a given point p;
4. the connection is *metric-compatible*: $\nabla_c g_{ab} = 0$.[28]

Condition 4 is of particular importance. It ensures that in (pseudo-)Riemannian geometry, the connection can be *defined* in terms of the metric,[29] rather than being independent of it.[30,31] Hence, there is only one fundamental mathematical object in Riemannian geometry, viz. g_{ab}; all the other quantities arise from it.[32]

Note that the connection is *not* a tensor. Whereas the most important property of a tensor is that if its components are zero/non-zero in one coordinate system, they are also zero/non-zero in *any* other coordinate system, this is not true for the connection. In particular, when changing from one coordinate system to another, the connection components transform in such a way that it is *always* possible to find a coordinate system in which the components of the connection are equal to zero at a given point.[33]

Now, back to our three interpretations. If the metric g_{ab}, the connection $\Gamma^a{}_{bc}$ and the Riemann tensor $R_{abc}{}^d$ *all* had geometric significance, but only, say, the

[25] I will use the so-called abstract index notation (cf. Wald 1984, Section 2.4). This means that whenever I use Latin indices, I denote a tensor as an *abstract* object. Writing Greek indices denotes the *components* of the tensor in question.

[26] If the connection is symmetric ($\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$), there is just one curvature tensor associated with the connection, if the connection is asymmetric ($\Gamma^a{}_{bc} \neq \Gamma^a{}_{cb}$), then it gives rise to *two* curvature tensors, both of which remain uniquely determined; cf. Goenner (2004, p. 17).

[27] This is the only condition that distinguishes pseudo-Riemannian geometry from Riemannian geometry: in the latter, the signature in four dimensions is $(++++)$.

[28] The covariant derivative ∇_c can be defined by the equation $\nabla_a T^b = \partial_a T^b + \Gamma^b{}_{ac} T^c$ where T^c is a vector.

[29] See Wald (1984, pp. 35–36) for a proof. If the condition does not hold, this gives rise to the so-called non-metricity tensor $Q_{ab}{}^c := g^{cd} \nabla_d g_{ab}$. In Weyl's theory from 1918, the first attempt to unify gravity and electromagnetism, the non-metricity tensor is assumed to have the special form $Q_{ijk} = Q_k g_{ij}$, where Q_k is then interpreted as corresponding to the electromagnetic vector potential. See Goenner (2004) for a summary of Weyl's theory and a comparison with other unified field theories.

[30] The connection of (pseudo-)Riemannian geometry is called the Levi-Civita connection; its components are given by terms containing only first derivatives of the metric.

[31] Furthermore, the condition ensures that the geodesics defined by the connection (via equation (1)) are the same as the geodesics of a local Lorentz frame, which is hence an inertial frame; see Misner et al. (1973, pp. 312–314).

[32] Even though it is generally not possible to deduce the metric and the connection from knowing the curvature tensor, it *is* possible to deduce a lot of information about the restrictions that the metric and connection have to fulfill if a specific curvature tensor is assumed—the latter places severe constraints on metric and connection; see Rendall et al. (1989) and references cited therein.

[33] The transformation that makes the components of the connection vanish will in general differ from point to point, i.e. it is not possible to 'transform away' the connection components throughout a neighborhood.

Riemann tensor had gravitational significance, then the geometric interpretation would have an argument in its favor. For we would have geometrical phenomena that were not connected to gravity at all, by being describable in terms of g_{ab} and $\Gamma^a{}_{bc}$ alone, whereas it would not be possible to have gravitational phenomena that are not connected to geometry, given that the Riemann tensor $R_{abc}{}^d$ has both geometrical and gravitational significance.

Similarly, if all the mathematical objects in the formalism had gravitational significance, but only some of them had geometrical significance, the field interpretation would be strengthened.

Finally, if every mathematical object in the formalism had *both* geometric and gravitational significance, then this would be an argument in favor of an egalitarian interpretation.

As established by Brown, the metric tensor g_{ab} of GR *does* have geometrical significance: because of the strong equivalence principle, a relation to rods and clocks is possible,[34] and I argued above that this is surely a sufficient condition for a mathematical object to have geometrical significance. Likewise, every other mathematical object in the formalism of GR has geometrical significance: while the metric can be linked to distances and durations between spacetime points, the connection defines which paths are inertial, i.e. geodesics, and the Riemann tensor represents the curvature of spacetime; all of them playing a role in accounting for geometrical phenomena like the ones described above.[35]

The question remains as to which mathematical objects have gravitational in addition to geometrical significance. This leads us to a related discussion: is it possible to speak of a physical object rightfully called 'the gravitational field', in quite the same way as we speak of 'the electromagnetic field'? If so, the question arises which of the mathematical objects in the formalism of GR *represents* the physical gravitational field. An answer to this question, or even just the endeavor to find one, is surely a promising way to find out which parts of the formalism have gravitational significance in the sense defined above.

4. WHAT IS A GRAVITATIONAL FIELD IN GR?

4.1 Candidate 1: the connection

The first candidate for a mathematical representative of the gravitational field is the connection $\Gamma^a{}_{bc}$. Janssen and Renn (2006) argue that seeing the connection components (Christoffel symbols) $\Gamma^\nu{}_{\mu\sigma}$ as representing the components of the

[34] See Chapter 9 of Brown (2005), where the strong equivalence principle (SEP) is defined (on p. 170) as a combination of two conditions, namely minimal coupling (= the laws of special relativity are valid 'in a sufficiently small region of spacetime') and universal coupling (= all non-gravitational interactions determine the same affine connection, i.e. pick out the same local inertial structure). Brown also argues (p. 160) that because of the dependence on the strong equivalence principle, "the 'chronogeometric', or 'chronometric', significance of $g_{\mu\nu}$ is not given a priori". This is surely true—but given that the SEP is part of the formalism of GR, the metric has chronometric significance *as a matter of fact*. We will see in Section 6 that also in rival theories of GR there is a case to be made about whether a given mathematical object has geometrical significance, i.e., using the terminology introduced above, whether it gives us a description and explanation of geometrical phenomena.

[35] Arguably, this looks as if the field interpretation has already lost the game. However, we will see in Section 5 that there are still variants of the field interpretation that remain available.

gravitational field (as doing so rather seeing than the derivatives of the metric $g_{\mu\nu,\sigma}$) was *the* realization that allowed Einstein to find the final field equations of GR.[36] It seems that Einstein did not change his mind on the issue—in a letter to von Laue from 1950, he writes:[37]

> It is true that in that case the R_{iklm} vanish, so that one could say: "There is no gravitational field present." However, what characterizes the existence of a gravitational field from the empirical standpoint is the non-vanishing of the Γ^l_{ik}, not the non-vanishing of the R_{iklm}. If one does not think intuitively in such a way, one cannot grasp why something like a curvature should have anything to do with gravitation. In any case, no reasonable person would have hit upon such a thing. The key for the understanding of the equality of inertial and gravitational mass is missing.

The second to last sentence recalls the context of discovery, but the last sentence seems to address a systematic rather than a historical point: Einstein claims that if the Riemann tensor were regarded as representing the gravitational field, there would be no theoretical link between inertia and gravity within the theory. What is meant here?

Let us have a look at the two versions of Einstein's famous elevator thought experiment.[38] In one version, it asserts that we cannot distinguish between a situation in which we are in an elevator, standing on the ground and being subject to a homogeneous gravitational force (or rather a homogeneous gravitational force *field*, since we feel the same force at every point in space), and a situation in which no gravitational field (nor the earth) is present, while we move with a uniform acceleration, thus being subject to a homogeneous inertial force (or rather a homogeneous inertial force *field*).

Pre-theoretically, we may regard the two situations as physically different: in one of them a gravitational force field is present, in the other we have an inertial force field.

Describing the second situation mathematically, we realize that inertial fields do not exist objectively, but that the phenomena associated with them in fact result from us being in a certain frame of reference. And since the situation where we regard ourselves as at rest and being subject to a homogeneous gravitational force field is empirically indistinguishable from this situation, it can likewise be associated with us being in a certain frame of reference rather than us being in an actually different situation.

[36] Janssen and Renn (2006, p. 839) point out that Einstein's Zurich notebook shows that he had already considered these field equations in 1912, but abandoned them in favor of the field equations he and Grossmann took as the core of the so-called *Entwurf* theory. In 1915, Einstein found his way back to what is now called the Einstein field equation. Janssen and Renn argue that the switch to regarding the connection components as representing the components of the gravitational field was the crucial step that convinced (and allowed!) him to do so. Compare also Renn and Sauer (2006, p. 161).

[37] Einstein to von Laue, September 12, 1950, Document 16-148 of the Einstein archives. Unfortunately, the letter is not yet available in the Collected Papers of Einstein (1987–); it is cited by Norton (1989, p. 39). As the context of the quote, Norton points out that "Laue had just pointed out that the Riemann–Christoffel curvature tensor vanishes in the context of the rotating disc argument."

[38] In many textbooks on GR, the elevator thought experiment(s) is used as a motivation for GR that is not looked at again once the full theory is available. However, it can be treated in the full theory, and as we will see, doing so is indeed quite instructive.

Indeed, both situations can be described by the connection $\Gamma^a{}_{bc}$ having certain components $\Gamma^\nu{}_{\mu\sigma}$ in a certain frame of reference, and thus there is no fact of the matter of whether a 'uniform inertial field' or a 'homogeneous gravitational field' is present *physically*. The connection gives us an account of what we originally thought of as two different force fields: in a given frame of reference, the inertial or gravitational field that we think we experience is accounted for by the components of the connection.

But there is *nothing* within GR that would justify us calling some components of the connection 'inertial components' and others 'gravitational components'. Hence, seeing the connection as the mathematical representative of gravity entails seeing the distinction between gravity and inertia as a matter of pre-GR terminology. It thus seems reasonable to adopt Ehlers' terminology of speaking of the connection $\Gamma^a{}_{bc}$ as the *gravitational-inertial field*,[39] which one can shorten to *'GI-field'*, analogous to EM-fields in Maxwell's theory.

Indeed, this unification is quite similar to how Maxwell unified electric fields and magnetic fields as electromagnetic fields. But it is not quite the same—the unification in GR is of a stronger kind. In electromagnetic theory (formulated in terms of 4-vectors), the electric and magnetic field are only aspects of the electromagnetic field, defined only with respect to a given frame of reference. But once a reference frame is chosen, there is still a fact of the matter whether a given component of the Faraday tensor is an electric or a magnetic component; for since there are electric, but no magnetic, monopoles that give rise to the electromagnetic field, an electric component $\vec{E}$ is a polar vector, whereas a magnetic component $\vec{B}$ is an axial vector. In GR, there is no similar asymmetry, and hence inertial components and gravitational components are indistinguishable from each other.[40]

Note that all this does not mean that we now have something like a gravitational-inertial *force* field—on the contrary! The connection does not only account for what we thought (!) of as homogeneous gravitational and inertial forces, it also determines which paths are to be regarded as force-free: it determines the geodesics of the spacetime structure. And in both situations described above, we move on geodesics, and hence are force-free according to the formalism.

In order to clarify this point, let us now come to the second version of the elevator thought experiment. It asserts that there is no way of distinguishing between being subject to no forces at all, and between freely falling while being subject to both gravitational and inertial forces. The reason is of course the equivalence of inertial mass and gravitational charge, and it is thus that in GR free-fall motions are defined to be the standard = inertial = force-free motions. They are therefore described by the geodesic equation:

$$\frac{d^2x^a}{d\lambda^2} + \Gamma^a{}_{bc}\frac{dx^b}{d\lambda}\frac{dx^c}{d\lambda} = 0 \tag{1}$$

[39] Ehlers (1973) switches between speaking of the connection $\Gamma^a{}_{bc}$ as accounting for the 'inertial-gravitational field' (p. 1) and as accounting for the 'gravitational-inertial field' (p. 20). Renn and Sauer (2006, p. 156) use a similar terminology when they describe Einstein's 'Pathways out of classical physics'.

[40] Maudlin (1996) claims that the unifications in electromagnetic theory and GR are both examples of 'perfect unification' (of electricity and magnetism on the one hand, inertia and gravity on the other hand). The above seems to show that instead in GR we find a stronger unification than in electromagnetism. But note that so far we have only tackled homogeneous gravitational and inertial fields—the general case is yet to be made.

where λ is an affine parameter. All this is only possible because the connection is not a tensor (cf. Section 3), and hence its components can always be 'transformed away' at a point, making the geodesic equation look like Newton's first law.

A connected, but additional, requirement is the existence of *local inertial frames*: again possible only because of the freedom in choosing the connection components. A local inertial frame is defined to exist at a point e if there exists some neighborhood of the point such that coordinates x^μ can be found in which, at e

$$g_{\mu\nu} = \eta_{\mu\nu}, \quad g_{\mu\nu,\sigma} = 0, \quad \Gamma^{\nu}{}_{\mu\sigma} = 0 \tag{2}$$

where $\eta_{\mu\nu}$ is the Minkowski metric. One could argue that seeing the connection $\Gamma^{a}{}_{bc}$ as representing a *physical* GI-field is synonymous with saying that GI-fields do not *really* exist, for how could something exist whose components can be 'transformed away', even if just locally? But as Giulini points out,[41] the vanishing of the components of the connection does not mean that the geometrical object 'connection' vanishes in some sense. Indeed, a connection is not a 'less good' geometrical object than a tensor, it is just a different kind of mathematical object: for a tensor, it is a non-trivial assertion whether its components vanish or do not vanish in a particular coordinate system. For a connection, the value '0' in a given coordinate system is just not a special point, because it does not tell us anything about the values the components can have in other coordinate systems—they can still be unequal to zero in such cases. For Giulini, this opens up the possibility of turning the tables and arguing that seeing the connection as the representative of the gravitational field does *not* mean that the gravitational field is 'never really there', but on the contrary that there really is no spacetime which is not endowed with a gravitational field. There is always a coordinate system in which the connection components are unequal to zero, and hence one is never justified in judging spacetime to be without a gravitational field. Likewise, we would never expect a spacetime without a geometry—a parallel that will become important later.

Furthermore, seeing the connection as the mathematical representative of the gravitational field makes the apparent problem of the energy-momentum of the gravitational field not being representable by a tensor appear in a different light. Just like the connection components, the components of the pseudo-tensor (which is not the same as a connection!) sometimes taken to represent the energy-momentum of the gravitational field can be made to vanish locally. Hence, one could say that it is indeed *to be expected*, rather than being a problem, that the energy-momentum associated with an entity that can be transformed away (locally), can itself be transformed away (locally) also.[42]

Even if one does not want to go so far as to say that the connection represents a *physical* gravitational-inertial field, it surely has gravitational significance in the sense defined in Section 3. For as we have seen above, it gives us an account of

[41] Giulini (2002), p. 24, in particular footnote 15, although his comments are of a rather brief character. The following thus partly rests on a private communication with Giulini on March 5, 2007: to whom my thanks.

[42] Cf. Bergmann (1976, p. 197). The point made remains valid even if the gravitational field is regarded as not being represented by the connection alone: a position advocated below. For independently of that, the components of the energy-momentum pseudo-tensor $t^{\nu}{}_{\mu}$ contain only *first* derivatives of the components of the metric. And since every first derivative of the metric can always be re-expressed by a combination of Christoffel Symbols (cf. Bergmann, 1976, p. 195), $t^{\nu}{}_{\mu}$ remains associated solely with the connection.

some of the most paradigmatic gravitational phenomena: for example the fact that things fall. And it represents the components of an apparent gravitational force in a given frame of reference.

Still, some have argued that the Riemann curvature tensor, rather than the connection, should be seen as representing the physical gravitational field. Indeed, seeing the connection as the sole representative of gravitational fields does not allow one to describe *inhomogeneous* gravitational fields—for that, we need the Riemann tensor. Even more importantly, we need the Riemann tensor in order to describe how the mass-energy-momentum of matter acts as a source for gravitational fields. This already suggests that the Riemann tensor has gravitational significance, but the questions of the *way* in which it has gravitational significance, and whether it can be rightfully seen as representing a physical gravitational field, are more difficult.

4.2 Candidate 2: the Riemann tensor

Synge writes:[43]

> [T]he first thing we have to get a feel of is the Riemann tensor, for it *is* the gravitational field—if it vanishes, and only then, there is no field.

The context of this quote is Synge's criticism of the principle of equivalence, of which he writes:

> Does it mean that the effects of a gravitational field are indistinguishable from an observer's acceleration? If so, it is false. In Einstein's theory, either there is a gravitational field or there is none according as the Riemann tensor does or does not vanish.[44]

But Synge does not distinguish between homogeneous and inhomogeneous gravitational fields. As we have seen in Section 4.1, *homogeneous* gravitational fields *are* indeed indistinguishable from *uniform* accelerations. This is exactly the point of Einstein's original principle of equivalence (Einstein, 1907). As Norton (1989) has pointed out, Einstein was well aware that the latter could not be extended to arbitrary gravitational fields.

Before I discuss Synge's view, I will briefly review how the Riemann tensor can be seen as being related to the inhomogeneity of an assumed physical gravitational field. We have already established (in Section 4.1) that freely-falling particles travel on geodesics. In a curved (4−)geometry, i.e. in a geometry where the Riemann tensor has non-vanishing components, the geodesics are generally non-parallel. Hence, freely-falling particles approach or diverge from each other, depending on whether the curvature is positive or negative. This behaviour is analogous to that which we would expect from particles moving in an inhomogeneous gravitational field present in (3−)space: they approach each other if the surface exerting the field has positive (2−)curvature and diverge from each other if the surface has negative

[43] Synge (1960, p. VIII), his emphasis.
[44] Synge (1960, p. IX).

(2−)curvature.[45] The measure of the approaching or diverging, which can be seen as the measure of the inhomogeneity of the gravitational field, is given by the equation of geodesic separation or geodesic deviation:[46]

$$\nabla_{\vec{u}}\nabla_{\vec{u}}V^a = R_{abc}{}^{d}U^bU^cV^d. \tag{3}$$

$R_{abc}{}^{d}$ is a tensor, and hence geodesic deviation is a coordinate-independent fact; it cannot be 'transformed away' like the components of the connection. Still, this does not make the gravitational field a force field: no matter if a particle moves in a homogeneous or in an inhomogeneous gravitational field (with no other fields being present), it always moves on geodesics, and hence is always force-free, even if it sees its own path approaching that of another particle. (Hence, talk of 'tidal *forces*' is very misleading in the context of GR.)

However, the consequence of seeing the Riemann tensor as the *sole* representative of the gravitational field is that it is not even possible to *interpret* the (homogeneous) force we feel when standing on the surface of the earth as a gravitational force, even if we go on and say that it is only apparently a force because our world-line is a geodesic. For Ghins and Budden (2001, pp. 38–39), this is sufficient to reject the Riemann tensor $R_{abc}{}^{d}$ as the mathematical representative of physical gravitational fields, in favor of the connection $\Gamma^a{}_{bc}$. However, they are a bit quick here: Synge could be quite happy with rejecting this distinction, claiming that only inhomogeneous gravitational fields are *real* fields anyway. Synge could argue that, after all, there are no perfectly homogeneous surfaces in nature, and hence no bodies which would actually exert a homogeneous gravitational field.

Granted, a body exerting a totally homogeneous field is an idealization. Not only would the body have to be perfectly flat, it would also have to be infinitely extended in order for the field not to become inhomogeneous around the edges. But do we really want to say that the more an actual body approaches this idealized situation, the more we lose the right to speak of it producing a real gravitational field? That would be like saying that a circle, if it were an absolutely perfect circle such as we do not find it in nature, should not be called 'a circle' since it would be too perfect a circle!

As already pointed out in Section 2.1, Einstein even regarded Minkowski spacetime *as it occurs within general relativity*, namely as a solution of the Einstein field equations,

$$G_{ab} := R_{ab} - \frac{1}{2}g_{ab}R = 8\pi G T_{ab} \tag{4}$$

as a special gravitational field. Indeed, the reason for this seems quite straightforward: it is asserted that a solution of the field equations can be used in order to describe a particular gravitational field.

Of course, the left-hand side of the Einstein equation (4) contains part of the Riemann tensor $R_{abc}{}^{d}$. The latter can be decomposed into terms containing the

[45] Note that this is generally true only if the surface exerting the field really is *just* a surface, i.e. if it does not extend in three dimensions. For general three-dimensional bodies, it is only true if the body has certain symmetries: a homogeneous sphere, for example, will exert the same form of field as if it was only the surface of a sphere—or a point mass.

[46] V^a is the separation vector, $\nabla_{\vec{u}}$ is the covariant derivative in the direction of the tangent vector $\vec{u}$, U^b the 4-velocity of one of the two particles.

Ricci tensor R_{ab} and the Ricci-scalar R on the one hand, and a term given by the Weyl tensor C_{abcd} on the other hand:[47]

$$R_{abcd} = C_{abcd} + \frac{2}{n-2}(g_{a[c}R_{d]b} - g_{b[c}R_{d]a}) - \frac{2}{(n-1)(n-2)}Rg_{a[c}g_{d]b} \tag{5}$$

At the same time, equation (5) is the *definition* of the Weyl tensor. In four dimensions, the Riemann tensor has twenty independent components, ten of which are 'delivered' by the terms containing Ricci tensor and scalar, the other ten by the Weyl tensor.[48]

Since the source term on the right-hand side of (4) is the energy-momentum tensor of matter T_{ab}, it is only natural to regard the Riemann tensor as representing the gravitational field that interacts with matter.[49]

A fuller account would now go on to describe the form of the constraints the Einstein field equations place on possible gravitational fields. By investigating the solution-space of the Einstein equation we would aim to find out what *kinds of fields* gravitational fields are. For example:

(i) Are they fully determined by the matter distribution? (No, only constrained.)
(ii) Do they give rise to transverse or longitudinal waves? (Both.)
(iii) Is there a superposition principle for gravitational fields? (No, the field equations are non-linear.)

Nobody can deny that the Riemann curvature tensor has gravitational significance—I am sure that not even the proponents of taking the connection as representing physical gravitational fields intended to do that. For we have seen that like the connection, the Riemann tensor gives us an account of paradigmatic gravitational phenomena: e.g., the phenomenon that things fall non-parallel, as described above.

4.3 Candidate 3: the metric—and all that it determines

The last two subsections have shown that both connection $\Gamma^a{}_{bc}$ and Riemann tensor $R_{abc}{}^d$ have gravitational significance in the sense defined in Section 3, while reviewing the discussion of whether one or the other should be regarded as the mathematical representative of physical gravitational fields.

However, the fact remains that there is only one *fundamental* mathematical object in the formalism of GR, the metric tensor field g_{ab}. If we insisted that only *one* mathematical object of the formalism should be called 'the gravitational field', it should be the metric g_{ab}, for all the other mathematical objects having gravitational significance derive from it. And of course, the metric has gravitational significance

[47] Cf., e.g., Wald (1984, p. 40) and Stephani (2004, p. 61). n is the dimension of spacetime, and the brackets stand for the process of antisymmetrization: $A_{[x}B_{y]} = A_xB_y - A_yB_x$.

[48] Cf. Stephani (2004, p. 61) or Bergmann (1976, pp. 172–174).

[49] Looking at the Einstein equation (4), one might be tempted to argue that it is the Einstein tensor G_{ab}, rather than the full Riemann tensor $R_{abc}{}^d$ which represents the gravitational field. But this would be shortsighted: the Weyl part of the Riemann tensor interacts with matter through the equation (derived from Einstein equation and Bianchi identity) $\nabla^a C_{abcd} = 8\pi G(\nabla_{[c}T_{d]b} + \frac{1}{3}g_{b[c}\nabla_{d]}T)$.

as well: it is what gives us the solutions to the Einstein field equations (4), thereby (together with the connection) being essential for describing phenomena like the bending of light by the sun. Likewise, in order to describe gravitational waves, we need to assume the metric to have a specific form.

Hence, the most sensible answer to the question 'What is a gravitational field in GR?' is not

(a) the connection $\Gamma^a{}_{bc}$ (Giulini, Ghins and Budden); nor
(b) the Riemann tensor $R_{abc}{}^d$ (Synge),

the answer is that the gravitational field is represented by the metric field g_{ab} and all the mathematical entities arising from it. I have only clarified the roles of the metric g_{ab}, the connection $\Gamma^a{}_{bc}$ and the Riemann tensor $R_{abc}{}^d$. But one may supplement the list by the Einstein tensor G_{ab}, the Ricci tensor R_{ab} and the Weyl tensor C_{abcd}. As already indicated in the last subsection, it is possible to extend the discussion to these entities, specifying further what *kind* of fields gravitational fields are. The Einstein tensor (and thereby the Ricci tensor and scalar) represents gravitational degrees of freedom that are *determined* by the matter distribution (via the Einstein equation), whereas the Weyl tensor encodes the degrees of freedom of gravitational fields that are not determined but only *constrained* by matter. Even the Bianchi identity might be seen as a close analogue to the second Maxwell equation, giving us constraints on the form that free gravitational waves can take.[50]

Looking back, the coordinate representations of the connection $\Gamma^a{}_{bc}$ and the Riemann tensor $R_{abc}{}^d$ already suggest that we see the metric as the *potential*,[51] the connection components as the *field strength*[52] and the Riemann tensor as the *rate of change* of the gravitational field (strength).[53] For the connection components are defined in terms of first derivatives of the metric, the components of the Riemann tensor in terms of the first and second derivatives of the metric.

But even if one does not like this language, or if one wants to abandon all talk of physical gravitational fields in a way that some proponents of the geometric interpretation might favor—claiming that there are no gravitational fields but only gravitational phenomena that are ultimately reduced to geometrical phenomena—what remains is that every mathematical object in the formalism of standard (vacuum) GR has both gravitational and geometrical significance.

[50] I thank Andrew Hodges for this last point.

[51] This term is used in various places to denote the metric, for example in Ehlers (1973, p. 21). Renn and Sauer (2006, p. 135) even claim that what they call the mental concept of a field is always related to the concept of a potential, the field arising by some differential operator acting on the potential. They make it very clear that for Einstein the metric was the gravitational potential in GR (p. 155). Note also that in the Newtonian limit of GR, in which the field equations are linearized by regarding the metric as just perturbatively differing from the Minkowski metric ($g_{\mu\nu} = \eta_{\mu\nu} + f_{\mu\nu}$ where $f_{\mu\nu} \ll \eta_{\mu\nu}$), the 00-component of the metric perturbation $f_{\mu\nu}$ takes over the role of the Newtonian gravitational potential in the Poisson equation.

[52] Remember the point made by Renn and Sauer quoted at the beginning of Section 4.1: seeing the connection components as representing the components of the physical gravitational field was the realization that allowed Einstein to find the field equations of GR.

[53] Goenner uses a similar terminology (Goenner, 2004, p. 27).

5. GEOMETRIC, FIELD AND EGALITARIAN INTERPRETATIONS REVISITED

In Section 3, I argued that the egalitarian interpretation would be strengthened if it turned out that every mathematical object (in particular the metric g_{ab}, the connection $\Gamma^a{}_{bc}$ and the Riemann tensor $R_{abc}{}^d$) in GR had both geometric and gravitational significance. Indeed, this has been shown in the previous section. Hence we could argue that we should neither state that gravity is a manifestation of spacetime geometry (geometric interpretation), nor that spacetime geometry is a manifestation of the behaviour of gravitational fields (field interpretation)—rather, it seems that the interpretation most faithful to the formalism is the egalitarian interpretation.

But as pointed out in Section 2, the three interpretations are families of positions rather than single positions. Thus, it is sensible to ask to what degree specific variants of the geometric and the field interpretation are compatible with all mathematical objects having both gravitational and geometrical significance, and hence whether an egalitarian interpretation is merely supported by the above reasoning rather than enforced.

It seems that at least some variants of the geometric interpretation can accommodate the pairing of gravitational and geometric significance within the formalism. A proponent could say: I believe that gravity is a manifestation of spacetime curvature. But this does not commit me to saying that it is *either* the curvature tensor *or* the connection that gives me an account of gravitational phenomena. I could say that for gravity to be there (or more precisely: for there to be phenomena that were previously explained by recourse to postulated gravitational fields), I need curvature to be there. But given that, I may need to use all the mathematical objects that the formalism offers in order to describe a given gravitational or geometrical phenomenon, I can well accept that in this sense all mathematical objects of the formalism have both gravitational and geometrical significance.[54]

This variant of the geometric interpretation argues that for gravity to be a *manifestation* of spacetime curvature, it need not be describable *solely* in terms of curvature—it only requires that gravity is there only if curvature is there. Another variant of the geometric interpretation might be even less restrictive, saying that gravity is a manifestation of geometry in general (rather than curvature in particular), and claim that all this means is that we cannot have gravitational phenomena without a link to spacetime geometry, but that we may well have geometrical phenomena which are not connected to gravity.

These two variants of the geometric interpretation do indeed point to a form of *weak egalitarianism* that is compatible with a geometric interpretation of the formalism: it contents itself with saying that all mathematical objects have both gravitational and geometrical significance, without concluding that therefore gravity and geometry should be conceptually identified. The proponent of the geometric interpretation would then say that all the relevant mathematical objects play a role in accounting for both gravitational and geometrical phenomena, but that not all geometrical phenomena have to be interpretable as gravitational phenomena.

54 This paragraph heavily relies on a discussion with Oliver Pooley, to whom my thanks.

Moderate egalitarianism, on the other hand, asserts that gravity and geometry stand in a real one-to-one correspondence, allowing us to reinterpret every gravitational phenomenon as a geometrical phenomenon *and* vice versa. It is incompatible with the versions of the geometric interpretation presented above; but yet another variant of the latter could accept moderate egalitarianism while claiming that nevertheless the geometric explanation of certain phenomena is in some sense more fundamental than the explanation in terms of gravitational fields.

Examples for reinterpretations motivated by accepting moderate egalitarianism are: (i) gravitational waves that are reinterpreted as waves of change in the geometric structure; (ii) the bending of light near the sun which can be interpreted as being due to the sun causing a certain geometric structure around it or the sun producing a gravitational field of a certain form which deflects the light beam; (iii) solutions of the Einstein equation that are normally regarded as corresponding to the universe having a certain geometric structure as solutions representing certain configurations of a gravitational field.

However, the points made in Section 4 only *enforce* a weakly egalitarian interpretation, whereas they are *compatible* with moderate egalitarianism as well. But given that the weak version makes the moderate version look very natural, and given that the moderate version has great heuristic value, it seems a sensible working hypothesis for future research.

What about the field interpretation? Just like the geometric interpretation, the field interpretation is compatible with weak egalitarianism. A proponent of the field interpretation could argue: It is quite alright for all the relevant mathematical objects in the formalism to have both gravitational and geometrical significance—although my position would be strengthened if some objects of the formalism had gravitational but no geometrical significance, it is not a disaster if this is not the case. My point is then just that the geometrical significance of the formalism *stems* from the way gravitational fields couple to matter fields; because the coupling is of a universal nature, the gravitational field *acquires* geometrical significance.

Of course, the proponent of the field interpretation would have to say what he means by 'gravitational field', for which he could resort to the discussion of Section 4.

Alternatively, he could argue that he never meant to interpret GR in its standard formulation, but rather turn to a formulation in which gravity is regarded as a spin-2 field h_{ab} defined on flat Minkowski spacetime η_{ab}. The idea goes back to a paper by Fierz and Pauli (1939), but has been improved significantly since then.[55] However, Robert Wald points out (although he has himself done much to improve the spin-2 approach (Wald, 1986)) that "the notion of the mass and spin of a field require the presence of a flat background metric η_{ab} which one has in the linear approximation but not in the full theory, so the statement that, in general relativity, gravity is treated as a massless spin-2 field is not one that can be given precise meaning outside the context of the linear approximation."[56]

[55] See the extensive 1995 foreword by Preskill and Thorne to Feynman's 1963 lectures on gravitation (Feynman, 1995) for a review of the history and literature on the subject.

[56] See Wald (1984, p. 76).

Indeed, one of the most important achievements of the spin-2 approach is surely the recovery of the full Einstein field equations; in doing so, a change of the dynamical variable from h_{ab} to $g_{ab}(\eta_{ab}, h_{ab})$ is performed in such a way that the theory no longer depends on the background metric η_{ab} (Wald, 1986). It can then be argued that the full metric g_{ab} is only an effective notion, and that what remains fundamental is spacetime geometry η_{ab} and gravitational field h_{ab} separately: the gravitational field would be what makes the *effective* metric of spacetime g_{ab} deviate from its *actual* metric η_{ab}, which is flat.[57] However, given that the spin-2 strategy leads to a recovery of the full Einstein equations in terms of g_{ab}, nothing prevents one from regarding it merely as a device of derivation, and from using the geometric interpretation once the Einstein equations are obtained—even Feynman, one of the main proponents of the spin-2 strategy, does so at times.[58]

This brings us back to the egalitarian idea. Indeed, although Feynman prefers a variant of the field interpretation, he writes the following:[59]

> It is one of the most peculiar aspects of the theory of gravitation, that it has both a field interpretation and a geometrical interpretation. Since these are truly two aspects of the same theory we might assume that the Venutian scientists, after developing their completed field theory of gravity, would have eventually discovered the geometrical point of view. [...] In any case, the fact is that a spin-two field has this geometrical interpretation. [...] [We want] to understand how gravity can be both geometry and field.

Even though the last sentence sounds like strong egalitarianism (cf. below), Feynman would surely rather find himself in the position I called moderate egalitarianism: the combination of gravitational and geometrical significance of the mathematical objects within GR is of such a kind that it is always possible to *switch* between a field perspective and a geometrical perspective. This still allows to *prefer* one perspective over the other, and with this caveat both Feynman and Rovelli seem to acknowledge moderate egalitarianism.[60]

Finally, there is the possibility to adopt *strong egalitarianism*. It is incompatible even with the sophisticated variants of geometric and field interpretation presented above: it is a different ontological position. It asserts that gravity and geometry do not just stand in a one-to-one correspondence, but that they are conceptually identified. Formulating it in terms of gravitational fields, one could say: according to strong egalitarianism, every gravitational field *is* a geometry of spacetime.

If moderate egalitarianism is adopted, then we have good reason to go one step further and opt for strong egalitarianism—what better explanation could there be for the possibility of switching back and forth between the field perspective and the geometrical perspective than that the gravitational field and the geometry of spacetime are actually one and the same 'thing'? But again, we are not forced to take this view.

[57] This perspective bears some resemblance to how Rosen sees his bimetric theory; cf. Section 6.2.

[58] Feynman still claims that the geometric interpretation is not necessary; see Feynman (1995, p. 113). But he does not give arguments for why the field interpretation is necessary.

[59] See Feynman (1995, p. 113).

[60] Cf. the quote by Rovelli in Section 2.1.

Of course, adopting moderate or strong egalitarianism does not mean that both a geometric and a field view on a given phenomenon are equally natural, or even equally valuable for every given phenomenon. But this may be a strength of these two versions of the egalitarian interpretation: they encourage us to reformulate a phenomenon we are used to think of as gravitational/geometrical in order to find new ways of tackling a given problem: they allow us to use the heuristics of both the geometric and the field interpretation.

The question is now whether the possibility to adopt any form of egalitarianism is peculiar to GR, or instead a generic feature of (modern) theories of gravity. In order to answer this question, I will have a look at two rival theories of GR.

6. JUST GR?

As argued in Section 4.3, the only fundamental non-matter field of GR is the metric g_{ab}. It has both geometrical and gravitational significance, and hence an egalitarian position is possible with respect to GR. Is this the case in other theories of gravitation?

In alternative theories of gravity, there is often more than one fundamental non-matter field, which can have scalar, vector or tensor character. When Will (1993) discusses the parametrized post-Newtonian (PPN) formalism as a basis for a theory of gravitation theories, he distinguishes only between a theory postulating one or more gravitational fields on the one hand, and it allowing for one or more matter (i.e. non-gravitational) fields on the other hand. However, many authors regard their theories as introducing *some* mathematical objects representing gravity, and *others* which are supposed to represent the (often nondynamical) structure of spacetime, in particular spacetime geometry. I will thus continue to distinguish between talk of the *fundamental* non-matter fields of a given theory, i.e. those fields that cannot be defined in terms of other (mathematical) fields, and between a fundamental field having gravitational and/or geometrical significance. Only if every fundamental non-matter field in a given theory has both geometrical and gravitational significance will an egalitarian interpretation of the theory be possible.

Note that from a field perspective, there are two ways of thinking of a theory that has more than one fundamental mathematical field with gravitational significance. We could say that according to such a theory, there are several kinds of gravitational fields that are represented by the different mathematical fields, or we could think of the different mathematical fields as representing different aspects of the one and only physical gravitational field.

6.1 Brans–Dicke Theory

Brans–Dicke theory was first proposed in 1961 (Brans and Dicke, 1961) in order to give a theory that is more in accord with Mach's principle than standard general relativity. The way in which Brans and Dicke aimed to realize this was by assuming that the gravitational (coupling) constant G is after all not a constant, but a

field that varies from spacetime point to spacetime point.[61] The varying gravitational 'constant' is represented by an additional scalar field ϕ. Thus, Brans and Dicke wanted to construct a theory with *two* fundamental non-matter fields with gravitational significance, the scalar field ϕ and the metric field g_{ab}.[62] One of the most important constraints in doing so was the desire to keep the geodesic equation of test particles unchanged as compared to standard GR: the scalar field was supposed to influence the behaviour of test particles only *indirectly*.[63] However, the standard Lagrangian for GR including matter is

$$L = \sqrt{-g}\,(R + 16\pi\kappa L_M) \tag{6}$$

where κ is the gravitational constant and L_M the matter Lagrangian. Simply substituting the constant κ by a scalar field ϕ would clearly change the geodesic equation for test particles. Hence, the following Lagrangian is chosen instead:[64]

$$L = \sqrt{-g}\left(\phi R - \frac{\omega}{\phi} g^{ab}\nabla_a\phi\nabla_b\phi + 16\pi L_M\right) \tag{7}$$

where ω is the dimensionless Brans–Dicke coupling constant and L_M a standard matter Lagrangian. Note that the scalar field ϕ indeed couples only indirectly to the matter fields by coupling to the Ricci scalar R and hence to the geometry, which *in turn* interacts with matter.[65] The second term of the total Lagrangian is the standard Lagrangian of a free scalar field.

Varying L first with respect to the metric g_{ab} and then with respect to the scalar field ϕ gives the *two* field equations of Brans–Dicke theory:[66]

$$R_{ij} - \frac{1}{2}g_{ij}R = \frac{8\pi}{\phi}T_{ij} + \frac{\omega}{\phi^2}\left(\phi_{;i}\phi_{;j} - \frac{1}{2}g_{ij}\phi_{;k}\phi^{;k}\right) + \frac{1}{\phi}(\phi_{;i;j} - g_{ij}\Box\phi) \tag{8}$$

and

$$\Box\phi = \frac{8\pi}{2\omega + 3}T^a_{\ a}, \tag{9}$$

where $\Box$ is the covariant Laplace operator $\Box\phi = \phi^{;a}_{\ ;a}$. Hence, we have one equation that looks like Einstein's equations with some extra terms, and one wave-equation

[61] Note that this is a much more general hypothesis than the more recent idea that the gravitational constant G may have changed its value during the evolution of the universe. Here, the change is just over time, whereas at any given time (under the restriction of the conventionality of simultaneity of course), G has the same value everywhere in space. In Brans–Dicke theory, a variation over time *and* space is allowed.

[62] Weinberg (1972, pp. 157–160) discusses the scalar field as a field *in addition* to the gravitational field g_{ab}. But he links the original gravitational constant to the *mean value* of the scalar field, thus allowing it to remain a constant after all, whereas in Brans' and Dicke's original theory the gravitational constant varies from point to point because of the scalar field.

[63] Brans (2005) recently gave a historical reconstruction of the development of Brans–Dicke theory.

[64] The Lagrangian of Brans–Dicke theory formally is a special case of the Lagrangian of Jordan's theory (Jordan, 1955, 1959). But as Brans and Dicke themselves emphasize, the physical interpretation is very different: Jordan predicted the expansion of the earth. For the even more general Lagrangian of an arbitrary so-called scalar-tensor theory of which both Jordan theory and Brans–Dicke theory are (mathematically) special cases, see Flanagan (2004, p. 3819) and references cited therein.

[65] See Weinstein (1996) for an analysis of different types of coupling in Lagrangian theories, alongside a concise discussion of Brans–Dicke theory and the role of the metric in the theory.

[66] I will from now on use the common shorthand to abbreviate a covariant derivative with a semicolon, while a partial derivative is represented by a comma.

for the scalar field ϕ. Metric g_{ij}, connection $\Gamma^{i}{}_{jk}$ and Riemann tensor $R_{ijk}{}^{l}$ have the same geometric and gravitational significance the corresponding entities have in GR, by similar reasoning as in the previous sections.

But what about the scalar field ϕ? Brans and Dicke (1961) write (p. 928):

> It is not a completely geometrical theory of gravitation, as [...] gravitational effects are in part geometrical [by being described by g_{ij}] and in part due to a scalar interaction.

The claim is thus that the scalar field has gravitational, but no geometrical, significance. Note that the lack of geometrical significance does not follow simply because ϕ is a *scalar* field. In Newton–Cartan theory, we have a scalar field τ which represents absolute time, and hence certainly has geometrical significance: because of τ, two space-like separated events can be objectively simultaneous. In Brans–Dicke theory, on the other hand, the scalar field ϕ does not help us predict any geometrical phenomena; it does not represent an aspect of spacetime structure.

But ϕ does play a role in predicting gravitational phenomena. Because of ϕ, the field equations of the theory allow for more solutions corresponding to gravitational waves than Einstein's theory. Furthermore, the theory predicts the so-called Dicke–Nordtvedt effect, a violation of the equality of inertial and gravitational mass for extended bodies due to composition-dependent couplings that make extended bodies fall in a non-universal way.[67] Were this effect observed, it should be regarded as a purely gravitational effect because the effect would amount to a violation of the weak principle of equivalence (cf. Section 3), and it is only because of the latter that we can associate gravity and geometry in the first place. So ϕ cannot be a field with geometrical significance. But why not regard it as just another matter field and the Dicke–Nordtvedt effect as an effect that arises when other kinds of matter couple to this particular form of matter?

The reason is that the scalar parts on the right-hand side of equation (8) can not be regarded as forming an energy-momentum tensor for the scalar field.[68] It is hence as reasonable to regard ϕ as a non-matter field as it is to regard the metric g_{ab} as a non-matter field in both GR and Brans–Dicke theory.

We can conclude that in Brans–Dicke theory it is *not* the case that all the fundamental non-matter fields have both geometrical and gravitational significance, contrary to what we found in GR (where the only fundamental non-matter field is the metric g_{ab}).[69]

[67] See Nordtvedt (1968) for a derivation of the effect, and Damour and Vokrouhlicky (1996) for an analysis of the tests connected to the Earth-Moon-Sun system. As Will points out (1993, p. 126), all predictions of Brans–Dicke theory differ from the predictions of GR at most by corrections of $O(\frac{1}{\omega})$. Most recent data demands that $\omega > 40.000$.

[68] See Santiago and Silbergleit (2000) for a detailed discussion of this point. The authors propose to introduce extra conditions in order to allow the definition of an energy-momentum tensor of the scalar field, arguing that thereby the Einstein frame (see the next footnote) becomes the natural representation of the theory. But for us, the important point is that the theory as it stands makes a clear distinction between matter fields (for which an energy-momentum tensor can be defined in a natural way) and gravitational fields (metric g_{ab} and scalar field ϕ), for which this is not possible.

[69] The above representation of Brans–Dicke theory is the so-called *Jordan frame* representation of the theory. Dicke (1962) has shown that it is possible to switch to the so-called *Einstein frame* of the theory by redefining ones measurement units (by help of a conformal transformation of the fundamental fields g_{ij} and ϕ) such that the rest masses of particles instead of the gravitational constant are position-dependent. Equation (8) then looks exactly like the Einstein equation of standard GR (4), the scalar field ϕ looks indeed like just another matter field, and test particles do not move on the geodesics of the newly defined metric. A further discussion could involve investigating in how far the gravitational and geometrical

6.2 Rosen's bimetric theory

Another rival theory of GR is Nathan Rosen's bimetric theory. The original motivation (Rosen, 1940a) of the theory was to reformulate GR in such a way that it would be possible to obtain a gravitational energy-momentum *tensor*, rather than the pseudo-tensor we have in standard GR. In order to accomplish this, Rosen proposes to base the theory on *two* metric tensors, the flat metric γ_{ab} and the curved metric g_{ab}, the latter of which we know from standard GR. In a follow-up paper, Rosen (1940b) argues that one could regard this as more than a mathematical reformulation, i.e. regard γ_{ab} as more than "a fiction introduced for mathematical convenience". Rosen claims that it is possible that[70]

> the metric tensor $\gamma_{\mu\nu}$ is given a real physical significance as describing the geometrical properties of space, which is therefore taken to be flat, whereas the tensor $g_{\mu\nu}$ is to be regarded as describing the gravitational field.[71]

Using our terminology from Section 3, Rosen seems to claim that in his theory the flat metric γ_{ab} has only geometrical significance, whereas the curved metric g_{ab} is supposed to have only gravitational significance, in contrast to GR.

The Lagrangian of the theory is given by:[72]

$$L = \frac{1}{64\pi}\left[\sqrt{-\gamma}\,\gamma^{ab}g^{cd}g^{ef}\left(g_{ce|a}g_{df|b} - \frac{1}{2}g_{cd|a}g_{ef|b}\right) + L_M\right] \tag{10}$$

where the vertical line "|" denotes the covariant derivative with respect to γ_{ab}.

The field equations derived from the Lagrangian are[73]

$$N_{ab} - \frac{1}{2}g_{ab}N = 8\pi\sigma G T_{ab} \tag{11}$$

where

$$N_{ab} = \frac{1}{2}\gamma^{cd}\left(g_{ab|cd} - g^{ef}g_{ae|c}g_{af|d}\right), \tag{12}$$

$$N = g^{ab}N_{ab} \quad \text{and} \quad \sigma = \sqrt{\frac{g}{\gamma}}. \tag{13}$$

Note that the similarity between Rosen's field equations (11) and Einstein's field equations (4) is of a rather superficial character: N_{ab} is not a Ricci tensor associated

significance of the mathematical objects changes in this representation of the theory. But this would raise a whole new topic, namely the status of conventions in establishing the physical significance of mathematical objects, and I will thus leave this for another occasion. For a review of the discussion of whether the two frames of Brans–Dicke theory should be regarded as equally physical, see Flanagan (2004), for a discussion of general scalar-tensor theories of gravity and Brans–Dicke theory in particular see Will (1993, p. 123), and Ni (1972) (both of which discuss the theories in the context of the PPN formalism).

70 Rosen has proposed 2 distinct bimetric theories, only the first of which is presented here. The first theory (Rosen, 1940a, 1940b, 1973) describes a theory with one flat and one dynamically curved metric, the second theory (Rosen, 1978, 1980a, 1980b) describes a theory with one metric of constant positive curvature and, like the earlier theory, one dynamically curved metric. Both theories face conceptually similar questions, in particular with respect to gravitational/geometrical significance of the two metrics; hence I will focus on the first theory.

71 Rosen speaks of $\gamma_{\mu\nu}$ describing the geometrical properties of space here. However, $\gamma_{\mu\nu}$ is a 4-metric, and thus surely meant to describe spacetime geometry.

72 See Rosen (1973, p. 441) and cf. Will (1993, p. 131).

73 g and γ stand for $\sqrt{det(-g_{ab})}$ and $\sqrt{det(-\gamma_{ab})}$, respectively.

with *one* metric like the Ricci tensor occurring in Einstein's equation, but a second-rank tensor of a different kind, being constructed out of covariant derivatives of one metric (g_{ab}) with respect to another metric (γ_{ab}).

What is the intuition behind Rosen's claims that γ_{ab} represents the geometry of spacetime and g_{ab} the gravitational field? Although Rosen himself has never discussed it, it seems helpful to look at his theory from the viewpoint of the parametrized post-Newtonian (PPN) formalism. The post-Newtonian limit of Rosen's theory differs from the one of GR:[74] the flat metric γ_{ab} establishes a preferred rest frame for the universe as a whole, and it has been shown that for any such theory, the gravitational constant G as measured by a Cavendish experiment on earth would depend on the earth's velocity relative to this preferred frame.[75] Furthermore, the particular kind of preferred frame we find in Rosen's theory allows for gravitational and electromagnetic waves to have different velocities as measured by an observer at rest in the preferred frame. To compare, in GR electromagnetic and gravitational waves have the same velocity, namely the velocity of light c.

The idea behind seeing γ_{ab} as representing geometry and g_{ab} as representing gravity can then be explicated by saying that because of having established a preferred rest frame, it seems sensible to regard γ_{ab} as representing the 'real' metric of spacetime. The gravitational field g_{ab} is then something that 'makes' things deviate from the actual geodesics of spacetime, the ones compatible with γ_{ab}, by 'making' them move on the geodesics of g_{ab}.

All this seems sensible, but it is after all very stipulative. As Rosen himself points out, the line element that is read by rods and clocks, and whose geodesics are followed by test particles,[76] is the one associated with the curved dynamical metric g_{ab}, just as in GR.[77]

Hence, by *our* criterion of what it means for a mathematical object to have geometrical significance, the metric g_{ab} *does* have geometrical significance in addition to gravitational significance.[78] Recall: in Section 3, it was argued that a relation to rods and clocks is at least sufficient for a mathematical object to have geometrical significance.

This does not deprive the flat metric γ_{ab} of its geometrical significance: the dependence of the results of a Cavendish experiment on a preferred frame of reference arguably qualifies as a geometrical phenomenon. However, γ_{ab} does not have gravitational significance: phenomena like gravitational redshift or the way things fall only depend on the curved metric g_{ab}. The flat metric γ_{ab} may place boundary conditions on what can be derived, and thus on the set of gravitational

74 See Will (1993, p. 117 and 131/132).

75 See Will and Nordtvedt (1972, p. 763).

76 The latter point is not emphasized by Rosen, who reformulates the geodesic equation in such a way that it looks different from standard GR, and in which he identifies a gravitational force term. But the important fact is that in his theory, just like in GR, the covariant derivative of the energy-momentum tensor *with respect to the curved metric* g_{ab} vanishes ($\nabla^b T_{ab} = 0$), and that hence particles follow the geodesics of this metric.

77 This is most clearly expressed in Rosen (1980b, p. 676). Although this text discusses Rosen's second bimetric theory, in which γ_{ab} is a metric of constant positive curvature, the same applies for the first metric theory (cf. Rosen, 1940b, p. 152).

78 That the curved metric g_{ab} of Rosen's theory has gravitational significance is very straightforward: gravitational phenomena like the bending of light by the sun and the existence of gravitational waves are derived in a way very analogous to GR; cf. Rosen (1940b, p. 152).

phenomena the theory can explain, but it does not play a role in describing the gravitational phenomena which the theory *does* cover.

Hence, just like for Brans–Dicke theory we can conclude that in Rosen's theory it is not the case that all the fundamental fields have both geometrical and gravitational significance. It thus seems reasonable that the possibility of an egalitarian interpretation—be it weak, moderate or strong—makes GR a rather peculiar member of the abstract 'space' of gravitational theories.

7. SUMMING UP AND CONCLUSION

I have presented three families of positions: the geometric, the field and the egalitarian interpretations of general relativity. The geometric interpretations claim that gravity is in some sense reducible to spacetime geometry, the field interpretations assert that spacetime geometry can be reduced to the behaviour of gravitational fields, and the egalitarian interpretations affirm that gravity and geometry stand in a one-to-one correspondence to each other.

I have proposed definitions of what it means for a given mathematical object to have geometrical and/or gravitational significance, in a way general enough for the definitions to not only apply to GR but to a wider class of theories. It was then shown that in GR all mathematical objects have both geometrical and gravitational significance, in particular the fundamental metric tensor g_{ab}. I have formulated three variants of the egalitarian interpretation, only the strongest of which is a certain alternative, the other two being compatible with at least some versions of the geometric and the field interpretation. Weak egalitarianism asserts that every mathematical object in the formalism has both geometrical and gravitational significance. Moderate egalitarianism claims that the correspondence between geometrical and gravitational significance is of such a kind that for any given phenomenon covered by the theory it is possible to switch back and forth between a geometric and a field picture of the phenomenon in question, thus allowing us to use the heuristics of both geometric and field interpretation. Finally, strong egalitarianism affirms that gravity on the one hand and spacetime geometry on the other hand are after all just two names for one and the same 'thing'.

I have argued that a weak egalitarian interpretation of GR is *enforced* by the formalism, but that there are also convincing reasons to adopt a moderate or even a strong egalitarian interpretation of GR.

I then turned to Brans–Dicke theory and Rosen's first bimetric theory, showing that the fundamental fields of the two theories do not all have both geometrical and gravitational significance, and that hence an egalitarian interpretation is not possible for these theories. But even if the scalar field ϕ in Brans–Dicke theory had geometric significance, and if the flat metric γ_{ab} in Rosen's theory had gravitational significance, only a weak egalitarian interpretation of the two theories would be possible. For it would still be the case that not all geometrical phenomena accounted for by Brans–Dicke theory involve ϕ, and neither would all gravitational phenomena described by Rosen's theory involve γ_{ab}. It would hence not be possible to opt for a moderate or strong egalitarian interpretation of either

of the two theories, for it would neither be possible to switch between a geometrical and a field perspective on any given phenomenon (moderate egalitarianism), let alone to conceptually identify gravity and the geometry of spacetime (strong egalitarianism).

Given that an egalitarian interpretation of Brans–Dicke theory on the one hand and Rosen's theory on the other hand does not seem possible, the question arises which version of geometric or field interpretation should adequately be adopted with respect to either of the two theories. Answering this question would demand a much closer look on the mathematical structure of the two theories, and a detailed analysis of the dynamics of each theory, but lack of space has prohibited me to deliver such an analysis.

Finally, it would be worthwhile to apply the categories established (the three families of interpretations, the notions of gravitational and/or geometrical significance of a given mathematical object) to a wider set of theories of gravity and spacetime structure, and to find out in how far they are useful in a general framework of a theory of gravitation theories.

ACKNOWLEDGEMENTS

I am very thankful to Jeremy Butterfield, Harvey Brown, Oliver Pooley, Eleanor Knox and Stephen Tiley. Each of them read different versions of this paper, and it surely would not be the same without their criticism and suggestions of improvement. Of course, this does not mean that they agree with everything, or that any remaining errors are to be blamed on anyone else but myself!

I am also grateful to audiences in Montreal, Oxford, Heidelberg, Bristol and Leeds for helpful and fascinating discussions.

REFERENCES

Anderson, J.L., 1967. Principles of Relativity Physics. Academic Press, Inc.

Belot, G., 1996. Why general relativity does need an interpretation. Philosophy of Science 63 (Supplement), 80–88. Proceedings of the 1996 Meeting of the Philosophy of Science Association.

Belot, G., 1998. Understanding electromagnetism. British Journal for Philosophy of Science 49, 531–555.

Bergmann, P.G., 1976. Introduction to the Theory of Relativity. Dover.

Brans, C., Dicke, R.H., 1961. Mach's principle and a relativistic theory of gravitation. Physical Review 124 (3), 925–935.

Brans, C.H., 2005. The roots of scalar-tensor theories: an approximate history. arXiv: gr-qc/0506063 v1.

Brown, H.R., 2005. Physical Relativity. Space-Time Structure from a Dynamical Perspective. Oxford University Press.

Damour, T., Vokrouhlicky, D., 1996. Equivalence principle and the moon. Physical Review D 53 (8).

Dicke, R.H., 1962. Mach's principle and invariance under transformation of units. Physical Review 125 (6), 2163–2167.

Earman, J., Glymour, C., Stachel, J. (Eds.), 1977. Foundations of Space-Time Theories. Minnesota Studies in the Philosophy of Science, vol. VIII. Minnesota Center for Philosophy of Science, University of Minnesota Press, Minneapolis.

Ehlers, J., 1973. Survey of General Relativity Theory. Astrophysics and Space Science Library, vol. 38. D. Reidel Publishing Company, pp. 1–123.

Ehlers, J., et al., 1972. The geometry of free fall and light propagation. In: O'Raifeartaigh, L. (Ed.), General Relativity. New York, Oxford.
Einstein, A., 1907. Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen. Jahrbuch der Radioaktivität und Elektronik 4, 411–462.
Einstein, A., 1916. Hamiltonsches Prinzip und allgemeine Relativitätstheorie. Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 1111–1116.
Einstein, A., 1917. Über die spezielle und die allgemeine Relativitätstheorie (Gemeinverständlich). Vieweg. English translation published as: On the Special and the General Theory of Relativity (a Popular Exposition). Methuen, 1960.
Einstein, A., 1918. Über Gravitationswellen. Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 154–167.
Einstein, A., 1949. Autobiographical Notes. The Library of Living Philosophers. Open Court Publishing. Page numbers refer to the reprint as a separate edition in 1979.
Einstein, A., 1987–. Collected Papers. Princeton University Press.
Eötvös, R.V., 1889. Über die Anziehung der Erde auf verschiedene Substanzen. Math. Naturw. Ber. aus Ungarn 8, 65–68.
Eötvös, R.V., Pekar, D., Fekete, E., 1922. Beiträge zum Gesetz der Proportionalität von Trägheit und Gravität. Annalen der Physik 68, 11–66.
Feynman, R.P., 1995. Feynman Lectures on Gravitation. Addison-Wesley Longman.
Fierz, M., Pauli, W., 1939. Relativistic wave equations for particles of arbitrary spin in an electromagnetic field. Proceedings of the Royal Society of London A 173, 211–232.
Flanagan, E.E., 2004. The conformal frame freedom in theories of gravitation. Class. Quantum Grav. 21, 3817–3829.
Ghins, M., Budden, T., 2001. The principle of equivalence. Studies in the History and Philosophy of Modern Physics 32 (1), 33–51.
Giulini, D., 2002. Das Problem der Trägheit. Philosophia Naturalis 39, 343–374.
Goenner, H.F.M., 2004. On the history of unified field theories. Living Rev. Relativity 7, 2. [Online Article]: cited [13.2.2004], http://www.livingreviews.org/lrr-2004-2.
Grünbaum, A., 1973. Geometrodynamics and ontology. The Journal of Philosophy 70 (21), 775–800.
Hartle, J.B., 2003. Gravity. An Introduction to Einstein's General Relativity. Addison-Wesley.
Janssen, M., Renn, J., 2006. Untying the knot: How Einstein found his way back to field equations discarded in the Zürich notebook. In: Renn, J. (Ed.), The Genesis of General Relativity, vol. 2. Springer.
Jordan, P., 1955. Schwerkraft und Weltall. Die Wissenschaft, vol. 107. Friedr. Vieweg und Sohn.
Jordan, P., 1959. Zum gegenwärtigen Stand der Diracschen Hypothesen. Zeitschrift für Physik 157, 112–121.
Maudlin, T., 1996. On the unification of physics. The Journal of Philosophy 93 (3), 129–144.
Misner, C.W., Thorne, K.S., Wheeler, J.A., 1973. Gravitation. Freeman.
Ni, Wei-Tou, 1972. Theoretic frameworks for testing relativistic gravity. IV. A compendium a metric theories of gravity and their post-Newtonian limits. The Astrophysical Journal 176, 769–796.
Nordtvedt, K. Jr., 1968. Equivalence principle for massive bodies. II. Theory. Physical Review 169 (5), 1017–1025.
Norton, J., 1989. What was Einstein's principle of equivalence? Einstein Studies 1.
Norton, J., 1992. Einstein, Nordström and the early demise of Lorentz-covariant, scalar theories of gravitation. Archive for History of Exact Sciences 45, 17–94.
Reichenbach, H., 1957. Philosophy of Space and Time. Courier Dover Publications.
Rendall, A.D., Hall, G.S., Kay, W., 1989. The curvature problem in general relativity. General Relativity and Gravitation 21 (5), 439–446.
Renn, J., Sauer, T., 2006. Pathways out of classical physics. Einstein's double strategy in his search for the gravitational field equations. In: Renn, J. (Ed.), The Genesis of General Relativity, vol. 1. Einstein's Zurich Notebook: Introduction ans Source Introduction and Source, Springer.
Rosen, N., 1940a. General relativity and flat space. I. Physical Review 57, 147–150.
Rosen, N., 1940b. General relativity and flat space. II. Physical Review 57, 150–153.
Rosen, N., 1973. A bi-metric theory of gravitation. General Relativity and Gravitation 4 (6), 435–447.
Rosen, N., 1978. Bimetric gravitation theory on a cosmological basis. General Relativity and Gravitation 9 (4), 339–351.

Rosen, N., 1980a. Bimetric general relativity and cosmology. General Relativity and Gravitation 12 (7), 493–510.
Rosen, N., 1980b. General relativity with a background metric. Foundations of Physics 10 (9/10), 673–704.
Rovelli, C., 1997. Halfway through the woods: contemporary research on space and time. In: Earman, J., Norton, J. (Eds.), The Cosmos of Science. University of Pittsburgh Press, Pittsburgh, pp. 180–223.
Santiago, D.I., Silbergleit, A.S., 2000. On the energy-momentum tensor of the scalar field in scalar-tensor theories of gravity. General Relativity and Gravitation 32, 565–581.
Stephani, H., 2004. Relativity: An Introduction to Special and General Relativity, 3rd edition. Cambridge University Press. Originally published in German as "Allgemeine Relativitätstheorie" in 1988. Page numbers refer to this version.
Synge, J.L., 1956. Relativity: The Special Theory. Noth-Holland Pub. Comp.
Synge, J.L., 1960. Relativity: The General Theory. Noth-Holland Pub. Comp., Amsterdam.
Wald, R.M., 1984. General Relativity. The University of Chicago Press, Chicago and London.
Wald, R.M., 1986. Spin-two fields and general covariance. Physical Review D 33 (12), 3613–3625.
Weinberg, S., 1972. Gravitation and Cosmology. John Wiley and Sons.
Weinstein, S., 1996. Strange couplings and space-time structure. Philosophy of Science Association 63 (Supplement).
Will, C.M., 1993. Theory and Experiment in Gravitational Physics. Cambridge University Press.
Will, C.M., Nordtvedt, K. Jr., 1972. Conservation laws and preferred frames in relativistic gravity. I. Preferred-frame theories and an extended ppn formalism. The Astrophysical Journal 177, 757–774.

FURTHER READING

Einstein, A., 1916. Näherunsweise Integration der Feldgleichungen der Gravitation. Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 688–696.

CHAPTER 6

Structural Aspects of Space-Time Singularities

Vincent Lam*

Abstract

We investigate the possible relevance of space-time singularities (within the theory of general relativity) for the debate about the nature of space-time. Standard attempts to describe space-time singularities in terms of local entities and local properties are discussed. It seems that space-time singularities possess some non-local or global aspects in the sense that they violate some basic aspects of (pre-)locality, which are inherent in the standard differential geometric representation of space-time. These possible non-local or global aspects of space-time underline the fact that the debate about the nature of space-time should not focus only on local aspects of space-time (such as space-time points) and should not be too dependent on one specific mathematical representation. In particular, we briefly discuss the possible relevance of the algebraic formulation of the theory of general relativity for the 'problem' of space-time singularities. Based on these considerations, a structural realist interpretation of space-time is proposed.

1. INTRODUCTION

Despite the invitation of John Earman to consider more carefully the question of space-time singularities, only a small part of the literature in space-time philosophy has been devoted to this foundational issue.[1] This chapter aims to take up this invitation and to carry out philosophical investigations about space-time singularities in the framework of the contemporary debate about the status and the nature of space-time. Indeed, taking into account the famous singularity theorems one

* Department of Philosophy, University of Lausanne, Switzerland

[1] Besides Earman (1995), some notable exceptions are Curiel (1999) and Mattingly (2001). The present chapter is an extension of Lam (2008).

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00006-5
© Elsevier BV
All rights reserved

may adopt two main positions with respect to space-time singularities and their generic character: first, they can be thought of as physically meaningless, only revealing that in these cases the theory of general relativity (GR) breaks down and must be superseded by another theory (like a future theory of quantum gravity (QG)).[2] Therefore, as such space-time singularities do not tell us anything physically relevant. Second, space-time singularities can be taken more 'seriously': they can well be considered as physically problematic but nevertheless relating to fundamental features of space-time. In this case, their careful study at the physical, mathematical and conceptual level may be helpful in order to understand the nature of space-time as described by GR. This chapter aims to investigate this line of thought. From this perspective, the question of space-time singularities is actually a fascinating one, which may be related at the same time to the question of the 'initial' state of our universe and to the question of the fundamental structure of space-time.

Roughly, the main question of this chapter is the following one: in a scientific realist perspective and assuming that the space-time singularities tell us something about the nature of space-time (again, this assumption is not evident), what do they tell us? The (tricky) problem of the very definition of space-time singularities is an essential part of the question.

2. SOME ASPECTS OF THE SINGULAR FEATURE OF SPACE-TIME

2.1 Extension and incompleteness

At the present state of our knowledge, it seems to be quite commonly accepted in the relevant physics literature that there is no satisfying general definition of a space-time singularity.[3] In other terms, the notion of a space-time singularity covers various distinct aspects that cannot be all captured in one single definition. We certainly do not pretend to review all these aspects here. We rather want to focus on the first two fundamental notions that are at the heart of most of the attempts to define space-time singularities.

The first is the notion of *extension* of a space-time Lorentz manifold (together with the interrelated notion of continuity and differentiability conditions).[4] The idea is to insure that what we count as singularities are not merely (regular) 'holes' or 'missing points' in our space-time Lorentz manifold that could be covered ('filled') by a 'bigger' but regular space-time Lorentz manifold with respect to some continuity and differentiability conditions (or C^k-conditions). These latter conditions (together with the notion of extension) are therefore essential for any characterization of space-time singularities. But, at this level, there are two major ambiguities that are part of the difficulties to define space-time singularities.

[2] They can be absent from our universe if one of the (necessary) conditions of the singularity theorems were violated, see Mattingly (2001). For a detailed physical discussion of these conditions, see Senovilla (1997).

[3] See for instance Wald (1984, 212).

[4] An extension of a space-time Lorentz manifold (M, g) is any space-time Lorentz manifold (M', g') of same dimension such that $g'|_{\varphi(M)} = \varphi^*(g)$, where φ^* is the 'carry along' map corresponding to the (imbedding) map $\varphi : M \to M'$ (that is, M and $\varphi(M)$ are diffeomorphic and $\varphi(M) \subset M'$ is a proper open submanifold of M'; (M', φ) is called an envelopment of M).

First, extensions are not unique and all possible extensions must be carefully considered in order to discard (regular) singularities that can be removed by a mere regular extension. Given some C^k-conditions, we will always consider maximal space-time Lorentz manifolds.[5] A space-time singularity will therefore be defined with respect to certain C^k-conditions (and indeed should be called a C^k-singularity; these conditions are often implicit and not always mentioned). This fact leads to the second difficulty: it is not clear what are exactly the necessary and sufficient continuity and differentiability conditions for a space-time Lorentz manifold to be physically meaningful.[6]

Strongly related with the idea of extension, the second essential notion in order to give an account of space-time singularities is the notion of *curve incompleteness*, which is the feature that is widely recognized as the most consensual characterization so far of space-time singularities.[7] Moreover, it is actually curve incompleteness that is predicted by the singularity theorems as the generic singular behaviour for a wide class of solutions.[8] The broad idea is that we should look at the behavior of physically relevant curves (namely geodesics and curves with a bounded acceleration) in the space-time Lorentz manifold for 'detecting' space-time singularities (which actually do not belong to the space-time Lorentz manifold): in particular, the idea is that an (inextendible) half-curve of finite length (with respect to a certain generalized affine parameter) may indicate the existence of a space-time singularity. The obvious intuition behind this idea is that, roughly, the (inextendible) curve has finite length because it 'meets' the singularity (it must be clear that this way of speaking is actually misleading in the sense that the 'meeting' does not happen in the space-time Lorentz manifold). Pictorially, anything moving along such an incomplete (non-spacelike) curve (like an incomplete geodesic or an incomplete curve with a bounded acceleration) would literally 'disappear' after a finite amount of proper time or after a finite amount of a generalized affine parameter (again, we must be very careful when using such pictures; for instance, the event of the 'disappearance' itself is not part of the space-time Lorentz manifold). In more formal terms, a (maximal) space-time Lorentz manifold is said to be *b*-complete if all inextendible C^1-half curves have infinite length as measured by the generalized affine parameter (it is *b*-incomplete otherwise).[9] The link with the initial intuition comes from the fact that it can be shown that *b*-completeness entails the completeness of geodesics and of curves with a bounded acceleration (but not vice versa).

[5] A space-time Lorentz manifold (M,g) is maximal with respect to some C^k-conditions if there is no extension (M',g') where the metric g' is C^k at the boundary $\partial' M$ of $\varphi(M)$ in M'.

[6] A possible guideline would be to require that these conditions secure that the fundamental laws of GR, that is, the Einstein field equations and the Bianchi identity, are well defined, see Earman (1995, §2.7).

[7] See for instance Wald (1984, §9.1).

[8] However, the notion of curve incompleteness does not encompass all aspects of space-time singularities, like for instance certain aspects linked with the singular behaviour of the curvature or with the violation of the cosmic censorship. In a general way, we will often speak of the 'singular behaviour of space-time'—or 'singular feature of space-time'—in the very broad sense of anything involving any of these singular aspects; more will be said about the appropriate terminology in Section 3.

[9] The generalized affine parameter u for a C^1-half-curve $\gamma(t)$ is defined by $u := \int_0^t (\sum_{\alpha=0}^3 (V^\alpha(t))^2)^{1/2}\, dt$, where $V(t') = V^\alpha(t') e_\alpha(p)$, $p = \gamma(t')$, is the tangent vector expressed in the parallel propagated orthonormal basis e_α.

2.2 Boundary

2.2.1 Attempts to 'localize' space-time singularities

The most widely accepted standard definition of a singular space-time is the following one: a (maximal) space-time (Lorentz manifold) is said to be singular if and only if it is b-incomplete. However, b-incompleteness only indirectly refers (if at all) to space-time singularities in the sense of localized singular parts of space-time (like space-time points where something 'goes wrong'). Space-time singularities are actually not part of the space-time Lorentz manifold (M,g) representing space-time (within GR) in the sense that they cannot be merely represented by points (or regions) of (M,g) where some physical quantity related to the space-time structure (like the scalar curvature for instance) goes to infinity.[10] It is in this sense that, strictly speaking, space-time singularities are not part of space-time (at least as described by classical GR). However, for various (physical as well as philosophical) reasons that will be discussed below, one would still want to describe them directly in terms of some local properties. In order to do that, one usually tries to 'attach' a kind of (singular) boundary ∂M to the manifold M, that is, a set of points (the 'missing points') to which some local properties can be ascribed. Mainly four different ways to attach a singular boundary to the space-time manifold have been developed. In this chapter, we will consider in some details only the two of them that can be considered as the less restrictive accounts of singularities in terms of boundary points, namely the b-boundary and the a-boundary. Furthermore, these two boundary constructions are representative of the kind of problems encountered when trying to ascribe space-time singularities with some local properties. The other two boundary constructions, the g-boundary and the c-boundary can be considered as more restrictive accounts (they are however interesting in their own right; the point is that they do not seem to bring anything additional for our discussion here).

2.2.2 b-boundary

The main idea of the b-boundary construction is to consider that the b-incomplete curves define (singular) boundary points (as their endpoints), which are then 'attached' to the space-time Lorentz manifold. Schmidt's procedure[11] provides a way to construct such a (singular) boundary (called b-boundary) using the equivalence between the b-completeness of the space-time Lorentz manifold (M,g) and the Cauchy completeness of the total space OM of the orthonormal frame bundle $\pi : OM \to M$.[12] Starting from a b-incomplete (M,g) (and therefore from a Cauchy incomplete OM), the idea is to construct the metric space completion $\overline{OM}$ and to extend the projection π and the action of the structure group $O(1,3)$ to $\overline{OM}$, so

[10] Space-time is represented within GR by (an equivalence class of pairs (M',g') diffeomorphic to) (M,g), where M is in general assumed to be a 'nice' (paracompact, connected, Hausdorff, oriented) 4-dimensional differentiable manifold and g is a C^k ($k \geqslant 2$ in general) Lorentz metric, solution of the Einstein field equations and defined everywhere on M. The link between active general covariance (or invariance under the active diffeomorphisms), which requires to take into account equivalence classes of pairs (M,g), and the singular structure of space-time will be discussed below.

[11] See Schmidt (1971).

[12] Unlike M, OM does admit a positive definite metric (actually determined by the Lorentz connection on M) and can therefore be considered as a metric space in a natural way. The important result is then that OM is Cauchy complete (with respect to the positive definite metric) if and only if (M,g) is b-complete (see Hawking and Ellis (1973, §8.3)).

to define the new base space as the quotient space $\overline{M} = \overline{OM}/O(1,3)$. But one can then express the new base space $\overline{M}$ as nothing else but the original base space together with a boundary, $\overline{M} = M \cup \partial M$,[13] so that this construction allows to define the so-called b-boundary $\partial M := \overline{M} \backslash M$. Indeed, within Schmidt's construction, every b-incomplete curve in (M, g) can be understood as determining a point in the b-boundary ∂M. At this stage, it seems legitimate to think about these boundary points as representing space-time singularities, in the sense of 'localized' entities (to which could then be ascribed 'local' properties). In order to translate (and to verify) this intuition of 'localization' in physical terms, it is necessary to endow the singular boundary with some differential or at least some topological structure. But this is where the problems show up (and the intuition of 'localization' breaks down). Indeed, it has been shown that the b-boundary of the closed FLRW solutions, which constitute part of the so-called 'standard model' in contemporary cosmology, consists of a single point that is not Hausdorff separated from points of the space-time Lorentz manifold M.[14] This is very problematic for two reasons: first, the closed FLRW solutions have two singularities, the one of which corresponding to the (singular) beginning of the universe and the other to its (singular) end. But within the b-boundary framework, these two singularities (the 'beginning' and the 'end' of the universe) are represented by one and the same boundary point, whose physical interpretation is then very difficult. Second, being not Hausdorff separated from points of M, this boundary point, which should represent the singularities, is actually 'arbitrarily close'[15] to the (regular) points of the space-time Lorentz manifold M. Again, it is very difficult to give physical meaning to such a behaviour since any (regular) points $p \in M$ has the singular boundary point (that represents the initial and final singularities in the FLRW solutions) in his (arbitrarily small) neighbourhood:[16] at least any (usual) sense of 'localization' of the singularities seems then to be lost—indeed one of the main motivations for attaching boundary points to the space-time Lorentz manifold is lost. Moreover, such bad topological behaviour has been shown to be a feature of all boundary constructions that share with the b-boundary construction certain natural (and rather weak) conditions.[17]

2.2.3 a-boundary

In order to overcome these difficulties, the a-boundary construction has been proposed as a somewhat different approach to the (local) attempts to characterize the space-time singularities in terms of boundary points.[18] It aims to truly capture the fundamental motivation of (most of) the boundary constructions, namely, the idea of 'missing points': the idea that, in order to be described in terms of local

[13] Since $OM = M/O(1,3)$ and since $O(1,3)$ 'preserves' Cauchy incompleteness (see Hawking and Ellis (1973, 283)); $\overline{M}$ is called the b-completion of M.

[14] See Bosshard (1976) and Johnson (1977).

[15] A topological space $\overline{M}$ is Hausdorff if $\forall p, q \in \overline{M}, p \neq q, \exists$ open sets $U, V \subset \overline{M}$ such that $p \in U$ and $q \in V$ and $U \cap V = \emptyset$; p and q are said to be Hausdorff separated. So, if two points of a topological space (like $\overline{M}$) are not Hausdorff separated, it is not possible for them to find two ('arbitrarily small') disjoint neighborhoods (open sets): it is in this topological sense that they can be considered as 'arbitrarily close'.

[16] The same problem arises in the case of the Schwarzschild solution.

[17] See Geroch et al. (1982).

[18] See Scott and Szekeres (1994).

properties, space-time singularities have to be considered as points in a 'bigger' manifold. More precisely, the motivation of the a-boundary construction is that singularities in a space-time Lorentz manifold have to be considered as points (or subsets) of the topological boundary of the (image of the) manifold with respect to an envelopment (such subsets are called boundary sets). The main problem of such considerations is that the possible envelopments of a given manifold (as well as the extensions, see Section 2.1) are far from unique. A possible way out is then to consider equivalence classes of boundary sets (with respect to different envelopments) under the following relevant equivalence relation (called the mutual covering relation): if B' is a boundary set of M with respect to an envelopment (M', φ')[19] and if B'' is a boundary set of M with respect to a second envelopment (M'', φ''), then B' and B'' are said to mutually cover each other if for every neighbourhood U' of B' in M', there exists an open neighbourhood U'' of B'' in M'' such that $\varphi' \circ {\varphi''}^{-1}(U'' \cap \varphi''(M)) \subset U'$ (B' covers B'') and vice versa (B'' covers B').[20] The a-boundary $B(M)$ of M is then defined as the set of such equivalence classes that have a singleton as a representative boundary set.

Although the a-boundary $B(M)$ itself is defined independently of any further geometric structure on M (such as a metric or a family of curves), it gains physical relevance for the very definition and study of the (essential) singularities only through selecting a family of curves (among a set of families of curves that satisfy certain properties): it is only with respect to a family of curves that one can identify the a-boundary points that lie at a finite 'distance' and therefore that constitute possible relevant (essential) singularities. So, once a family C of curves is fixed, the a-boundary points can be classified into various classes. In particular, the class of 'real' (essential) singularities (called the abstract singular boundary) is defined as the set of a-boundary points (called the abstract singularities) whose representative boundary point is the limit point of some incomplete curve from C and cannot be covered by a boundary point that is not the limit point of some incomplete curve from C.[21] As with other boundary constructions, one could be tempted to think of these abstract singularities as representing space-time singularities in terms of localized entities (with possible local properties). However, if some topological properties can actually be defined for the abstract singular boundary,[22] there are fundamental difficulties with such an interpretation.[23] Indeed, an abstract singularity (a 'point' of the abstract singular boundary) is an equivalence class of boundary sets under the mutually covering relation. Apart from the representative singleton, these boundary sets are in general not singletons and are moreover not even necessarily connected, so that, in this framework, any interpretation of a space-time singularity as a 'pointlike' space-time entity to

[19] $B' \equiv \overline{\varphi'(M)} - \varphi'(M)$.

[20] That is, if for every neighbourhood U'' of B'' in M'', there exists an open neighbourhood U' of B' in M' such that $\varphi'' \circ {\varphi'}^{-1}(U' \cap \varphi'(M)) \subset U''$. Roughly, the idea is that two boundary sets are equivalent when one cannot approach one without approaching the other (see Scott and Szekeres (1994)).

[21] It is important to notice that this characterization of essential singularities is invariant under the equivalence relation of mutual covering and therefore is well-defined for a-boundary points.

[22] For instance, a-boundary points are compact in the sense that every set of the equivalence class that defines the a-boundary point is compact.

[23] A deeper analysis of the interpretational difficulties of the a-boundary has been carried out in Curiel (1999).

which could be ascribed (local) properties seems very difficult—again, one of the main motivations of the boundary approach to space-time singularities fails.[24]

Moreover, it is also important to notice that the *a*-boundary approach cannot account for the (possible) singular behaviour of compact space-time regions: for instance, a space-time Lorentz manifold where a *b*-incomplete curve is (partially) contained ('imprisoned') in a compact region will (automatically) be considered as non-singular from the *a*-boundary point of view (although it is *b*-incomplete).[25] Without discussing whether or not compact space-time Lorentz manifolds (or space-time Lorentz manifold containing compact regions) are physically relevant, we emphasize that it shows that the *a*-boundary approach is not equivalent to the *b*-incompleteness approach. Indeed, since a compact manifold cannot be embedded as a proper submanifold of another manifold of same dimension, this shows that *b*-incompleteness is not equivalent to the idea of 'missing points' (or of 'holes') that actually belong to an extended manifold in which the *b*-incomplete manifold can be embedded. More fundamentally, this shows that, in general, *b*-incompleteness is not equivalent to any (usual) notion of 'localization'[26] and to the extent that *b*-incompleteness indicates the presence of space-time singularities or, in other (more neutral) terms, a singular behaviour of the space-time structure (what is generally recognized), it seems therefore that this latter cannot be properly characterized in terms of local entities or in terms of local properties. We now discuss some possible physical as well as philosophical meanings of this failure.

3. NON-LOCAL ASPECTS OF THE SINGULAR FEATURE OF SPACE-TIME

3.1 Physical and philosophical background

As we have seen already (see Section 2.2), the singular behaviour of space-time (like the *b*-incompleteness predicted by the singularity theorems for a very wide class of solutions) cannot be naively described by space-time Lorentz manifold points (or regions) where something goes wrong (where the curvature 'blows up' for instance). This is actually intimately related to the dynamical nature of the space-time structure as described by GR: space-time singularities are indeed singularities of the space-time structure itself and there is no a priori fixed structure or entity with respect to which the space-time singularities could be defined.[27] From the point of view of standard GR, a space-time singularity is not described as an entity localized in space-time.

[24] Moreover, the very physical significance of the covering relation for (singular) boundary sets is not clear in so far as an abstract singularity can cover non-singular boundary sets (see (Curiel (1999, 135)); indeed, it remains an open question whether or not every abstract singularity always covers a pure singular boundary set (which is a singular boundary set that does not cover any non-singular boundary sets, see the conclusion of Scott and Szekeres (1994)).

[25] Every compact space-time Lorentz manifold is non-singular within the *a*-boundary framework (see Theorem 50 in Scott and Szekeres (1994)), for a compact manifold M cannot be embedded as a proper submanifold of a manifold M' of same dimension (see Hawking and Ellis (1973, 289)).

[26] Within the *b*-boundary construction, *b*-incompleteness contained in a compact region ('imprisoned incompleteness') can actually be characterized by non-Hausdorff behaviour of the *b*-boundary (see Hawking and Ellis (1973, 289)) and as we have seen already in Section 2.2.2, this latter topological property makes any intuitive sense of 'localization' very difficult.

[27] See Earman (1995, 28). These difficulties to meaningfully define space-time singularities are strongly related to the background independence of GR.

From the physical point of view, the main motivation for the various localization attempts to describe the singular behaviour of space-time is the possibility to do local physics about the singular behaviour of space-time, that is, to characterize it in terms of local properties (and in terms of 'singularities') and to study the space-time structure 'near' or even 'at' the singularities. Indeed, this move is part of the usual methodology of physics: in general, physicists proceed first by doing local physics, that is, by analyzing local physical systems and by trying to isolate them from the rest of the world; global considerations are then made out of these local analysis. However, things are getting more complicated within GR, since the Einstein field equations determine the very structure of space-time, locally as well as globally.[28]

Besides this fundamental methodological motivation, it is possible to see a metaphysical but no less fundamental motivation behind the localization attempts of the singular behaviour of space-time. Indeed, it is commonly believed that non-quantum physics, like (classical) GR, can be understood within—or even supports—the traditional metaphysical position according to which the world is fully described by the distribution of fundamental physical intrinsic properties at points—space-time points or point-sized bits of matter—that are connected by spatio-temporal relations. In accord with the standard view, we define intrinsic properties to be those that an object has irrespective of whether or not there are other contingent objects; in brief, having or lacking an intrinsic property is independent of accompaniment or loneliness.[29] We will often refer to this conception as the 'atomistic' conception or worldview:[30] indeed, it considers the world as made of 'atoms' in the philosophical sense of 'building blocks' possessing intrinsic properties and existing therefore independently of one another (even if they are linked by spatio-temporal relations). This traditional atomistic conception of the world is expressed in contemporary metaphysics within Lewis' famous thesis of Humean supervenience.[31] It seems to constitute the metaphysical background for the various attempts to localize the singular behaviour of space-time with the help of boundary points. The idea is to associate the notion of an event—and therefore of a (boundary) point—to the description of the singular behaviour of space-time, to which local (and possibly intrinsic)[32] properties could then be ascribed, since the very notion of event makes reference in general to a (highly) localized occurrence in space-time.

But we have seen that the characterization of the singular behaviour in terms of a boundary attached to the space-time Lorentz manifold (although interesting for various purposes) does not allow one in general to localize the singular behaviour in any satisfactory and meaningful sense, that is, it does not allow to express the singular behaviour in terms of meaningful local entities and properties, so that it can be qualified as 'non-local' in some sense. In order to discuss the meaning and

[28] About the subtle relationship between local and global aspects of space-time within GR, see Demaret et al. (1997).

[29] For the standard definition of intrinsic properties, see for instance Langton and Lewis (1998).

[30] It also corresponds to the 'pointillisme' against which Jeremy Butterfield argues within classical physics, see Butterfield (2006).

[31] See Lewis (1986, ix–x).

[32] The notions of 'intrinsic' and 'local' are not equivalent, for instance see Butterfield (2006, §2.1.2); for the precise definition of 'local' as used here, see Section 3.2 below.

the possible implications of these 'non-local' aspects, we need now to be precise about which aspects of 'locality' are violated by the singular behaviour of space-time.

3.2 Locality within the standard geometric representation of space-time

The notion of 'local' or of 'locality' within physics is indeed a tricky and difficult issue that comprises many different aspects, involving notions like finite speed of action, causation and determinism.[33] But for our discussion, we only need to focus on the basic aspects of locality, which are indeed captured—inherent—in the standard differential geometric representation of physical reality (in which GR is usually defined for instance) in terms of a Lorentz manifold (M, g), on which tensor fields are defined. Such a representation provides a description of physical reality that has built into it "a significant amount of locality" (or of "pre-locality").[34] This encoding has two main aspects.

First, some fundamental (pre-)locality is captured in the standard differential geometric representation through the (obvious) geometric fact that any tensor field on a space-time region R can be meaningfully restricted to subregions of R and in particular to (arbitrarily small neighbourhoods of) points of R, where it defines the corresponding tensors. Conversely, these restrictions to subregions of R can be considered as subsystems (ultimately tensors—or tensor field properties—determined at space-time points) that entirely determine the considered (non-gravitational) tensor field (which can therefore be considered as a set of such subsystems).[35]

The second aspect of the (pre-)locality encoding lies in the fact that the standard differential geometric representation precisely encodes the common ('local') conception of space-time—space-time as a set of points with certain relations and properties. Within this framework, *local* entities and properties are then naturally defined as the ones that can be associated with (and determined at) a space-time point and its (arbitrarily small) neighbourhood. In accord with the common sense, this definition also encompasses any extended entity that can be meaningfully associated with a space-time point in a standard differential geometric model (when looking at different scales for instance). It is obvious that there is no need for a space-time property to hold only at one single space-time point (and its arbitrarily small neighbourhood) in order to be local: a property of a space-time region R that holds in arbitrarily small neighbourhoods of every point of R is also defined as local.[36] The important point for our discussion here lies in the fact that, in order

[33] See Earman (1987) for a detailed discussion of the various aspects of locality within physics.

[34] See Earman (1987, 453–454): Earman discusses various different (eleven!) 'locality principles', among them the one discussed here and requiring roughly that the considered theory could be meaningfully expressed within the standard differential geometric formalism, that is, in terms of tensor fields over a smooth differentiable manifold. Since this demand seems mandatory in order to meaningfully express the other (standard) 'locality principles', Earman calls it "pre-locality".

[35] Together with some (pre-)locality, some basic aspects of separability are involved here; see Earman (1987, 453): "[...] the description of physical reality in terms of local geometric fields [...] demands semantic separability and semantic localizability (it implies that, given the state throughout M, it is meaningful to speak of the restriction of the state to proper subsets of M, and, conversely, the state throughout M can be seen as the 'sum' of the states at all points of M)."

[36] Curiel (1999, 137–139) offers a formalisation of this definition of a local property of a differentiable manifold, which curve incompleteness does not satisfy.

to be physically meaningful, such a definition of local entities and properties requires that a (sufficiently strong) topological separation assumption, namely the Hausdorff separation condition, holds among the space-time points, so that distinct local properties can be associated with (determined at) distinct space-time points and their distinct (arbitrarily small) neighbourhood, no matter how close to one another they are. In other terms, any meaningful attribution of local properties to space-time points requires that any two space-time points (no matter how close they are) can be attributed a disjoint (arbitrarily small) neighbourhood (more precisely, a disjoint open set), i.e. that they are Hausdorff separated. The Hausdorff assumption, which lies therefore at the heart of the definition of local entities and properties, is in general part of the standard differential geometric representation of space-time (see Section 2.2); in this sense, it constitutes a basic aspect of locality (as discussed here) that is encoded—inherent—in the standard differential geometric representation of physical reality.

3.3 Non-local aspects: some implications

It is precisely this basic aspect of locality that is violated by the singular behaviour of space-time in some physically important cases: the singular boundary points that should represent the space-time singularities do not (necessarily) fulfill the Hausdorff separation condition with the space-time Lorentz manifold points that represent space-time points. Therefore, they cannot be qualified as local in the above defined sense. This is what is meant when it is said that the singular behaviour is non-local and that it cannot be described in terms of meaningful local entities and properties. Within this framework we can more generally say that the singular behaviour of space-time (as described by standard GR), be it considered from the point of view of curve incompleteness or from the boundary point of view, may display (topological) properties that prevent any generic characterization referring to a local entity, as in the case of a b-incomplete compact space-time region (imprisoned b-incompleteness) or of a non-Hausdorff boundary.

Let us note the possibility to require the singular behaviour of space-time to be described in terms of boundary points, therefore not considering imprisoned b-incompleteness as indicating a singular feature of space-time. In particular, this implies to reject b-incompleteness as a sufficient criterion for a singular behaviour. Such a (strong) requirement is made by the proponents of the idea of 'missing points', as in the a-boundary construction (see Section 2.2.3). It should be clear that this position towards space-time singularities does not (necessarily) avoid the violation of locality discussed above, since the boundary points representing space-time singularities may (still) not be Hausdorff separated from the space-time Lorentz manifold. Moreover, depending on the boundary construction we choose, the local entity corresponding to a space-time singularity is not always clearly and uniquely defined: for instance, in the case of the a-boundary, as we stressed in Section 2.2.3, it is not always clear how the constructed boundary point can be interpreted as a local entity (and how it should be interpreted in general), even in some relatively simple cases (see footnote 24). In other terms, the very physical significance of the implementation of the 'missing point' requirement through the

boundary constructions seems in general not to be well established.[37] One way towards getting some physical significance would be to secure locality, that is, to further require the singular feature of space-time to be described by meaningful local entities, namely topologically well-behaved boundary points. But it should be clear that the physical as well as philosophical price to pay for such a requirement is rather high: physically, this means considering numerous *b*-incomplete space-times (sometimes involving explicit unphysical behaviour) as non-singular. From the philosophical point of view, it amounts to a priori imposing an understanding of the singular behaviour of space-time in local terms for a priori metaphysical reasons, namely to secure a priori a fundamental aspect of the traditional atomistic conception of the world.

However, if one wants to avoid a priori metaphysical speculations, one should develop a metaphysical conception of space-time that is grounded on our current best physical theories about space-time, like GR (and not the other way round). Within this framework, it seems that the singular behaviour of space-time is best understood as a non-local (or global) feature of space-time, since, as we have discussed above, there is no clear and necessary link between the singular behaviour of space-time and the existence of any particular (meaningful) local entities like space-time (boundary) points or local properties instantiated at particular space-time (boundary) points. In other terms, this singular feature of space-time seems to be an irreducible non-local feature of space-time in the sense that it is not based on the existence of—and cannot be reduced to—any particular local entities or properties. So the conception of the singular behaviour of space-time as involving space-time singularities understood as local entities seems to be actually misleading—as the very talk of 'space-time singularities' is indeed. Therefore, it seems more accurate to speak of 'singular feature of space-time' in a sense that is not committed to the existence of any local entity or property. As counter-intuitive at first sight as it may be, such an understanding has been put forward by some philosophers of physics as well as by some physicists working in GR.[38] However, its implications for the metaphysical conception of space-time (and for the debate between substantivalism and relationalism) have not been very much discussed yet. We propose here some elements of reflection on this question.[39]

[37] Curiel (1999, 135–136) raises several objections against the criterion of 'missing points' for space-time singularities and proposes in particular a rather simple argument against this criterion. He argues that, as a criterion for a singular feature of space-time, curve incompleteness is "logically prior" to (the addition of) 'missing points', since the boundary constructions rely in general on the notion of curve incompleteness (except, perhaps, for the *c*-boundary construction). Moreover, in many (physically important) cases, it is difficult to see what these added 'missing points' (whose structure can indeed be rather complicated and/or whose (topological) properties can be unphysical in a straightforward sense) bring to the physical understanding of the singular feature of space-time. Therefore some kind of parsimony principle applied to physics suggests that we should merely dispense with them.

[38] In particular, it is especially emphasized in Curiel (1999, 139): "There is no a priori reason to suspect that the existence of an incomplete curve, a global phenomenon, could be tied in any natural or reasonable way to the existence of a particular point in an extended manifold." See also, for instance, Dorato (1998, 340): "In fact, 'being singular' is, like 'being orientable', a *global* property of spacetimes: as such, *unlike* ordinary properties, *it cannot be instantiated somewhere and somewhen*, [...]" and Earman (1996, 630): "[...] determining whether singularities actually occur [...] involves [...] determining whether or not spacetime has some large-scale or even global properties." Among physicists, this understanding is made explicit for instance in Geroch and Horowitz (1979, 269): "Perhaps, for example, singular behavior is to be understood, not by means of local singular points, but rather as a more global property of the entire spacetime."

[39] Again, as underlined in the introduction, such a discussion may be relevant only if one makes the assumption that the singular feature of space-time have something to say about the nature of space-time; otherwise, it may constitute an instance of "futile metaphysical analysis" (Dorato, 1998, 340).

The notions of local properties and local entities as defined above are closely linked to the concept of a point and to the notion of intrinsic properties, although the two notions of locality and of intrinsicality are clearly distinct. An intrinsic property (instantiated at a space-time point or region) is necessarily a local property (but not vice versa). Therefore, because of its non-local character, the singular feature of space-time cannot be tied to some intrinsic properties instantiated at some particular space-time points or regions. More precisely, there are two related points here: first, the singular feature of space-time can be understood as a kind of property of the entire space-time in the sense that it does not supervene on the existence of a (set of) particular space-time (boundary) point(s) or on the instantiation of some intrinsic (or relational) properties by such points. Second, the singular feature of space-time does not (globally) supervene on the entire distribution of intrinsic properties at space-time points either. In this perspective, it bears some analogy with some non-local aspects of the gravitational energy and, in a certain sense, with some global topological features of space-time.[40] These two points make clear that the non-local character of the singular feature of space-time does not merely amount to the widely recognized non-supervenience of space-time relations on intrinsic properties of the corresponding space-time points.[41] Whereas a particular space-time relation needs to be instantiated between particular space-time points (it is non-supervenient only in the sense of the second point), what we want to stress here is that space-time may possess some fundamental features that are actually independent of the existence of any particular space-time points or regions. Space-time as a whole may have irreducible fundamental properties that cannot be reduced to and do not depend on the existence of local or even intrinsic properties instantiated at space-time points (or point-sized bits of matter), so that any atomistic conception may not be able to account for such fundamental properties of space-time. Therefore, such non-local features of space-time may challenge the received atomistic view of space-time as a set of points possessing some intrinsic properties together with some space-time relations (like within Lewis' thesis of Humean supervenience).[42] So, it seems that not only quantum physics, but also classical general relativistic physics may threaten this traditional metaphysical conception of the world.[43]

3.4 Singular feature of space-time and the debate between substantivalism and relationalism: structural aspects

The non-local character of the singular feature of space-time (if taken 'seriously'), and in general other possible non-local (global) properties of space-time, like the gravitational energy, strongly suggest that we should not restrict our ontological considerations about space-time only to local entities and properties. Indeed,

[40] Recall the quotations in footnote 38.

[41] See Cleland (1984).

[42] To include merely the non-local features of space-time in the supervenience basis would be a rather ad hoc solution; moreover, it would fundamentally modify the atomistic worldview in terms of fundamental independently existing entities (the 'atoms').

[43] Butterfield (2006) has recently argued that also classical mechanics excludes this atomistic conception about space-time and the world, which he calls 'pointillisme'.

any ontological position with respect to space-time—substantivalist and relationalist ones—needs to accommodate the irreducible non-local (global) properties of space-time. In other terms, both substantivalist and relationalist ontologies have to account for the non-atomistic aspects of space-time. So, insofar as they challenge the traditional atomistic worldview, the irreducible non-local (global) aspects of space-time should prevent us from putting too much ontological weight on local and/or intrinsic properties as well as on local entities, like space-time points or pointlike bits of matter.[44] Any ontological conception of space-time, be it substantivalist or relationalist, that seeks to account for these aspects should not give priority to local entities over the global structure in which they are embedded. In particular, substantivalism about space-time should not be (necessarily) committed to a substantivalist view about space-time points in the strong sense of (independent) entities possessing intrinsic properties.

The irreducible non-local (global) properties of space-time call for a global (and holistic) conception of space-time, where space-time can be meaningfully considered as an irreducible whole, to which can be ascribed fundamental non-local and global properties that do not supervene on and cannot be reduced to local states of affairs. In other terms, parts of this 'space-time whole'—space-time points or regions for instance—cannot be meaningfully considered as independent of one another (in some sense, the non-local (global) properties bind together the parts). In particular, such needs can be met within a structural conception of space-time—space-time structuralism or, in an explicit scientific realist perspective, space-time structural realism—according to which space-time is understood as a (purely relational) physical structure in the (broad) sense of a complex network of relations that are not supervenient on independent individual entities possessing intrinsic properties (such that the structure itself is not supervenient on independent individual entities). We consider here space-time structuralism or space-time structural realism as a metaphysical thesis about space-time in contrast to a mere epistemological thesis about the limits of our knowledge (again, in an explicit scientific realist perspective, space-time structural realism is an ontological thesis in the sense that it is a claim about what there is in the world, namely structures). Space-time structuralism naturally emphasizes global aspects of space-time in the sense that, within this framework, space-time is considered as an irreducible whole, that is, as a non-supervenient structure, to which can be ascribed irreducible (non-supervenient) non-local and global properties. So, space-time as a whole does have an 'internal' structure, which can be represented by various different geometrical and algebraic structures, and space-time structuralism, as a metaphysical claim about space-time, is not dependent on a particular representation of the space-time structure (see Section 4). Therefore, at this stage, the broad metaphysical slogan of space-time structuralism—'space-time is a structure'—remains neutral with respect to the kind of relata and relations of the space-time structure. The fundamental point of this structural framework is that relata and relations (whatever they are) are on the same ontological level, so that,

[44] The main claim here does not side with any position in the debates between substantivalism and relationalism and about the relationship between space-time and (energy-)matter, but for simplicity we mainly use the space-time points talk.

in some sense, relata, relations and in definitive the whole structure they are part of are (ontologically) 'given at once'. A straightforward consequence that is fundamental for our concern about the singular feature of space-time is that local states of affairs cannot be considered as ontologically prior to global aspects.

Recently, a more specific structural conception of space-time has been developed as a consequence of the common understanding of the GR-principle of active general covariance (or of invariance under active diffeomorphisms) and the related hole argument (we can also invoke here the related principle of background independence). Indeed, due to this common understanding (involving equivalence classes and the so-called Leibniz equivalence), a wide range of philosophers of physics and physicists agree on the fact that, within GR, space-time points cannot be physically individuated (and therefore 'localized'), possessing intrinsic properties for instance, independently of the space-time relations as represented by the metric.[45] Indeed, within the standard geometric framework, such an understanding of active general covariance (and of background independence) seems to vindicate a structural conception of space-time in the sense of a network of non-supervenient space-time relations among non-intrinsically characterized space-time points. Many of the most widely accepted conceptions of space-time—either labelled 'substantivalist' or 'relationalist'—that take into account active general covariance, the hole argument and background independence seem to adhere to this structural understanding.[46] To the extent that it is explicitly committed to space-time points (with structural identity), this latter differs therefore from the broader structural conception of space-time suggested by the non-local (global) aspects of space-time (see the discussion below).[47]

So, it seems that, if the substantivalist and relationalist ontologies (whatever they are) adhere to a (non-atomistic) ontology of structures, they can then naturally account for the irreducible non-local (global) properties of space-time. As we have seen, within this structural framework, the singular feature of space-time can be more appropriately understood as a property of the space-time structure conceived as a whole (and therefore trivially involving non-local (global) aspects). However, it should be clear that, if the traditional debate between substantivalism and relationalism is considered within the (broader) framework of the issue of the relationship between space-time and (energy-)matter, as we claim it should,[48] then, the shift to an ontology of structures does not constitute a tertium quid in the debate: indeed, this structural move does not provide a definite position with respect to the relationship between the space-time and the (energy-)matter structures.

[45] See for instance Dorato (2000), Rovelli (2004, Ch. 2), the contributions in Rickles et al. (2006), and Esfeld and Lam (2008).

[46] In particular, it seems to be the case for the recent 'sophisticated substantivalist' and 'non-reductive relationalist' positions, see the contribution of Dean Rickles and Steven French in Rickles et al. (2006).

[47] We are grateful to an anonymous referee and to Michael Esfeld for highlighting that point.

[48] It allows to define precisely the substantivalist and relationalist positions: relationalism is defined in the reductive sense as the reduction of space-time to (non-spatio-temporal) properties and relations of matter (Leibnizean relationalism); substantivalism is considered to be the position according to which either space-time is an independently existing entity that has its own properties, which are not reducible to the properties and relations of matter (Newtonian substantivalism) or space-time and matter are ontologically identical and form the same substantival entity (Cartesian-Spinozean substantivalism), see Esfeld and Lam (2008, §4).

The debate about the nature of space-time sometimes revolves around the issue of the existence of the space-time points. As we have seen above, the irreducible non-local (global) aspects of space-time suggest a kind of sceptical attitude towards space-time points and their possible intrinsic properties, at least as basic elements of the ontology. Such a sceptical attitude is also fundamentally entailed by the broad structural conception of space-time, where emphasis is put on structures rather than on independent local entities (indeed, this emphasis allows the very account of non-local (global) aspects of space-time). However, depending on the version of space-time structuralism under consideration, there are important differences with respect to the ontological status of the space-time points. According to the moderate version, the space-time points are considered as being on the same ontological footing as the space-time relations and structure: according to this version, in the same way that a concrete spatio-temporal relation only exists between space-time points, its relata, these latter only exist in virtue of their standing in relation to other space-time points (relational properties), therefore not possessing any intrinsic properties, which could ground their independent existence (concrete relational properties ground their identity in exactly the same way as intrinsic properties would; in cases where they fail to do so, numerical distinction or diversity can be taken as a primitive fact).[49] On the other hand, the radical version of structural realism, which is known as 'ontic structural realism', either considers the existence of space-time points as ontologically secondary in the sense of being derived from the (ontologically primary) space-time relations and structure (leading to a kind of bundle theory of identity for the space-time points) or merely denies their existence (the original and most radical position: there are only relations and no relata).[50] All these versions of structural realism about space-time are mainly grounded on active general covariance and background independence. The structural conception of space-time that is suggested by the non-local and global aspects of space-time involves a broader notion of space-time structure than in the moderate version for instance, since it is not specifically committed to space-time points. However, it does not face the conceptual difficulties of the most radical version (what is a concrete relation without relata?), since it does not claim that space-time is 'relations all the way down'. But to what kind of space-time structure exactly is it committed to? Ultimately, as we will discuss below, this is an empirical question (see Section 5). At this stage, our claim here is only that the non-local and global properties of space-time can be meaningfully accounted for within a broad structural conception of space-time—the important thing being that the space-time 'structure' (as a whole) and the space-time 'constituents' (whatever they are) are on the same ontological level (in the above discussed sense).

[49] Such moderate structural realism—sometimes also referred to as the 'no priority view'—has been recently elaborated and proposed as a coherent structural metaphysics of space-time as described by GR in Esfeld and Lam (2008); see also the contribution of Oliver Pooley in Rickles et al. (2006). This 'no priority view' is analogous to the structural conception suggested by the non-local (global) properties of space-time, apart from the explicit commitment to space-time points. Such a commitment may result from a too strong dependence on a particular representation of the space-time structure, namely the standard geometrical one (see Section 5).

[50] See French and Ladyman (2003); for a recent discussion of the relationship between relations and relata within structural realism, see the contribution of John Stachel in Rickles et al. (2006) (however, John Stachel omits the 'no priority view' of the moderate version presented above).

Indeed, even if it seems at first sight that, in a scientific realist perspective, we are ontologically committed to space-time points (realism towards physical fields seems to require realism towards space-time points (or regions)), it may be the case that the moral of the 'space-time singularities problem' is that the very concept of a space-time point is challenged at the fundamental level. According to such a radical approach to the question, the singular feature of space-time would reveal the fundamental non-local (global) and pointless nature of space-time. Insofar as some amount of locality is inherent in the standard differential geometric representation (see Section 3.2), space-time would need to be described within another mathematical framework. This latter could be algebraic.

4. ALGEBRAIC APPROACH TO THE SINGULAR FEATURE OF SPACE-TIME

4.1 Algebraic formulation of GR

If the philosophical analysis of the singular feature of space-time is able to shed some new light on the possible nature of space-time as described by GR (as we have tried to show above), one should not lose sight of the fact that the traditional description of the singular behaviour of space-time is physically unsatisfactory in the sense that it may involve the divergence of relevant physical quantities (at a finite distance) like the scalar curvature.[51] Since the singular feature of space-time is directly connected to fundamental issues about the nature of space-time, like the issue of the 'initial' state of our universe—and of space-time itself indeed—improving its description and its understanding amounts to gaining knowledge about the nature of space-time. As we said already, many physicists believe that such an improvement will (only) ultimately come from a (complete) QG theory. Within this chapter, we want to investigate the singular feature of space-time from the classical GR point of view and we now want to consider 'improvements' or rather theoretical developments (generalizations) already at the classical level. From the discussion above, the (irreducible) non-local aspects of space-time and in particular the singular behaviour of space-time can be understood as suggesting a challenge to the worldview according to which local entities and in particular space-time points are fundamental. Or suggesting that such a worldview is actually induced by and deeply entrenched in one specific (though usual) mathematical framework, namely the standard differential geometric representation of space-time within GR. There are two interrelated aspects here: first, the (rather speculative) need for theoretical developments of GR that do away at the fundamental level with local entities such as space-time points. Second, the need for an understanding of space-time that does not rely so heavily (only) on one specific mathematical formulation. With respect to these two aspects, the algebraic formulation of GR may be of some interest.

The mathematical roots of the algebraic representation of space-time within GR are to be found in the full equivalence of, on the one hand, the usual (geometric) definition of a smooth differentiable manifold M in terms of a set of points

[51] However this does not entail that GR is either false or incomplete, see Earman (1996).

with a topology and a smooth differential structure (compatible atlases) with, on the other hand, the alternative (algebraic) definition of M using only the algebraic structure of the (commutative) ring $C^\infty(M)$ formed by the set of the smooth real functions on M (under pointwise addition and multiplication; indeed $C^\infty(M)$ is an algebra). For instance, the existence of points of M is equivalent to the existence of maximal ideals of $C^\infty(M)$.[52] Furthermore, tensorial calculus and equations on M can be defined and carried out in purely algebraic terms, namely, starting uniquely from the algebraic structure of $C^\infty(M)$. The set of contravariant vector fields on M can be understood (and defined or 'built') in purely algebraic terms as the module $\mathcal{D}(M)$ over $C^\infty(M)$ constituted by the derivations on $C^\infty(M)$. Then, in this perspective, the dual module $\mathcal{D}(M)^*$ defines the set of covariant vector fields on M and a metric on M can be defined as a (symmetric) isomorphism $g : \mathcal{D}(M) \to \mathcal{D}(M)^*$. Tensor fields can be understood as multilinear maps on the product space formed by copies of $\mathcal{D}(M)$ and $\mathcal{D}(M)^*$ (the number of copies is given by the type of the tensor field). In an analogous way, a (compatible or metrical) covariant derivative ∇, its associated Riemann tensor R and the (usual) contraction operation on tensor fields (which allows to define the Ricci tensor for instance) can be defined in purely algebraic terms. So, the Einstein field equations and their solutions, which represent the different possible space-times, can be defined in purely algebraic terms from $C^\infty(M)$.[53] In a general way, one can then define an Einstein algebra $\mathcal{A}$ as a linear algebra (with no reference to the smooth differentiable manifold M) that allows the above (algebraic) constructions (Lorentz metric, compatible or metrical covariant derivative) and such that the (algebraic) Einstein field equations are satisfied. The important point is that the algebraic structure of $\mathcal{A}$ can be considered as primary in exactly the same way in which Lorentz manifold points or sets of Lorentz manifold points, representing space-time points or regions, may be considered as primary from the standard geometric point of view and the space-time Lorentz manifold (M, g), solution of the Einstein field equations, can be considered as being derived from this algebraic structure. Therefore, there is an emphasis on global aspects of space-time in the sense that manifold points (which usually represent space-time points within the standard geometric framework) are not postulated from the beginning as (irreducible) basic elements and therefore (pre-)locality may not be considered as fundamental either.

However, from the mathematical and physical point of view, this algebraic formulation of GR in terms of $C^\infty(M)$ is completely equivalent to the standard geometric formulation of GR, so that it can be argued that this alternative algebraic formulation is nothing but a mere renaming of the usual standard geometric entities (for instance renaming of "points of the manifold M" with "maximal ideals of the algebra $C^\infty(M)$").[54] Now, the important thing is that the above mentioned

[52] A maximal ideal of a commutative algebra $\mathcal{A}$ is the largest proper subset of—indeed a subgroup of the additive group of—$\mathcal{A}$ closed under multiplication by any element of $\mathcal{A}$. The maximal ideal of $C^\infty(M)$ that corresponds to a point $p \in M$ is the set of all vanishing functions at p.

[53] The original formulation of GR in algebraic terms is due to Geroch (1972), which, together with Heller (1992), we referred to for the detailed mathematical constructions. Recently Bain (2006) also presents and discusses the algebraic formulation of GR: we very much agree with the line of the paper and we try in our discussion below to further deepen the implications of the algebraic formulation of GR (in particular with respect to the singular feature of space-time).

[54] There is a bijection between the models of GR in terms of the algebra $C^\infty(M)$ and the models of GR in terms of the Lorentz manifold (M, g). This is the main reason for the claim that, as such, the shift alone to this alternative algebraic

emphasis on global aspects of space-time can be generalized within the algebraic formulation of GR in a way that may be relevant for dealing with the singular feature of space-time.

4.2 Generalization of the algebraic formulation of GR and the singular feature of space-time

Indeed, the algebraic formulation of GR can be generalized in many ways and at different levels. We want neither to review all these possible generalization moves nor to enter into too many technical details here. Rather, we want to highlight some aspects that are directly related to the question of the singular feature of space-time.[55]

First, it should be noted that (M,g), a solution of the Einstein field equations in the standard geometric formulation, induces (through its related algebra $C^\infty(M)$) a mathematical (sub)representation $\kappa : \mathcal{A} \to C^\infty(M)$—a subrepresentation of a Gelfand representation indeed—of an Einstein algebra $\mathcal{A}$. But $\mathcal{A}$ can be represented by other (sub)representations, leading to functional algebras C (called 'differential structure') more general than $C^\infty(M)$ and to spaces (called 'differential spaces') more general than smooth differentiable manifolds. Indeed, such differential spaces can be used to investigate the singular feature of space-time and in particular its non-local aspects:[56] within the framework of the b-completion $\overline{M} = M \cup \partial M$ of a b-incomplete space-time Lorentz manifold (M,g) (see Section 2.2.2)—singular solution of the Einstein field equations—it is possible to prolong the differential structure $\mathcal{C}$ underlying M to the differential structure $\overline{\mathcal{C}}$ underlying $\overline{M}$ such that $\overline{\mathcal{C}}_M = \mathcal{C}$, where $\overline{\mathcal{C}}_M$ is the restriction of $\overline{\mathcal{C}}$ to M. In the singular cases where the b-boundary is not Hausdorff separated from the space-time Lorentz manifold M, like in the important physical cases of the closed FLRW and Schwarzschild solutions, it has been shown that only constant functions can be prolonged to the b-completion $\overline{M}$. This underlines the non-local feature of the singular behaviour of space-time, since constant functions can be considered as non-local in the sense that they do not distinguish points.

So, we see that this generalized algebraic formulation of GR further highlights the non-local aspects of space-time and in particular those linked with the singular feature. In this move towards non-local aspects, the possibly 'ultimate' step is then the replacement of commutative functional (Einstein) algebras with non-commutative (Einstein) algebras and their associated non-commutative spaces for the description of (singular) space-time. A non-commutative space is (completely) non-local in the sense that its underlying non-commutative (Einstein) algebra has (in general) no maximal ideals, which are the algebraic structures that encode the concept of manifold point: indeed, in this sense, non-commutative spaces are

formulation of GR does not elude the hole argument, which can indeed be translated in algebraic terms; for a recent discussion of the implications of the algebraic formulation of GR on the hole argument, see Bain (2003) and references therein.

[55] We briefly discuss in particular the approach of Michael Heller and coworkers, see for instance Heller and Sasin (1999)—to which we refer for the technical details—for a brief review of some aspects of their work linked with the singular feature of space-time.

[56] This investigation can be done in the even more general framework of sheaf theory.

pointless spaces. Such an algebra can be completed to a C^*-algebra, which admits a standard representation in terms of an algebra of bounded operators on the relevant Hilbert space,[57] so that the global concept of a state of the non-commutative Einstein algebra can be defined (roughly, a positive normed linear functional on the algebra). Without entering into too many technical details, space-time is then considered at the fundamental level as a (global) physical system that can occupy, as a whole, various physical states, which do not encode any distinction between singular and non-singular states of affairs. Indeed, it seems that such distinctions vanish in the same way that, within this non-commutative framework, any notion of (pre-)locality and of local entities disappears at the fundamental level. From this point of view, the singular feature of space-time reveals, at the (less fundamental) level of the differential geometric framework of standard GR, a fundamental aspect of space-time, namely its fundamental non-local and pointless nature.[58] Obviously, if sound, this aspect of the singular feature of space-time would have deep consequences for our understanding of space-time, some of which we want to briefly address in the last section and to link with the previous part of the discussion (Section 3.4).

5. SPACE-TIME AS A STRUCTURE: SOME PERSPECTIVES AND OPEN ISSUES

As we have seen in Section 3.4, the debate about the nature of space-time has to deal with (but is not reduced to) the question of the existence of the space-time points (or regions). As far as this question is concerned, the non-commutative generalizations of the algebraic formulation of GR provides a clear answer, at least in the case of the pointless spaces. But what does it entail for the very nature of space-time? Does it involve, as we could be tempted to think at first sight, that since there are no space-time points within this framework, space-time itself simply does not exist? We think not. As we have discussed already, a substantival position about space-time is not (necessarily) committed to a substantival view about space-time points. Moreover, from the broader perspective of the relationship between space-time and (energy-)matter, pointless or non-commutative spaces are not committed to any position.

The rejection of space-time points (or regions) and of any local entities indeed constitutes the 'ultimate' step against any atomistic worldview. As we have seen already, this move may suggest a structural conception of space-time (see Section 3.4). Indeed, space-time structuralism may provide a positive account of such a generalized algebraic description of space-time: within this framework, the space-time structure is characterized by the algebraic structure of a (non-commutative) Einstein (C^*-)algebra $\mathcal{A}$ (via its states for instance), that is, a network

[57] Indeed, to some extent, the whole move can be understood as being motivated by the analogy with quantum mechanics, where the set of all observables is a C^*-algebra: the idea is then that such a move does not only help to cope with the singular feature of space-time but may also be useful to describe possible quantum gravitational effects.

[58] This fundamental aspect gives some meaning to the intuition that the fundamental nature of space-time is to be found in two 'directions', namely, 'backward in time' and 'deeper in space', so that the question of the initial state of the universe (and of space-time) is fundamentally linked to questions about the nature of space-time, see Heller (2001).

of algebraic relations, which stem from the formal definition of a (C^*-)algebra and from the (algebraic) constraints imposed by the Einstein field equations, among the abstract elements of $\mathcal{A}$.[59] Such an algebraic generalization of GR promotes a global understanding of space-time that seems to be strongly related to the structuralist's emphasis on non-atomistic structural aspects of space-time. Of course, at this stage, these theoretical generalizations of GR remain rather speculative (although "by no means esoteric")[60] and need empirical anchorage. Whether or not space-time at the fundamental level really is non-local and pointless is an empirical question—as well as the question to the structuralist about the kind of structure (the kind of relations, relata, ...) space-time instantiates in the world.

Besides these fundamental empirical questions it raises, our discussion of the singular feature of space-time underlines the fact that our interpretative framework for the debate about the nature of space-time should free itself from the traditional atomistic and local (pointlike) conception of the world and of space-time in particular (such a conception seems to be induced to some extent by the standard differential geometric formulation of GR). Indeed, this speaks in favour of a structural conception of space-time. Furthermore, the singular feature of space-time constitutes a case where an alternative formulation, namely the algebraic one—even if completely equivalent to the standard differential geometric formulation—provides fruitful insights that allow further theoretical developments (which, of course, need to be experimentally confirmed).[61] Therefore, it seems that such an algebraic formulation deserves to play a role in the interpretative issues about space-time—at least to the same extent as the standard differential geometric formulation does.[62]

REFERENCES

Bain, J., 2003. Einstein algebras and the hole argument. Philosophy of Science 70, 1073–1085.

Bain, J., 2006. Spacetime structuralism. In: Dieks, D. (Ed.), The Ontology of Spacetime. Amsterdam, Elsevier.

Bosshard, B., 1976. On the *b*-boundary of the closed Friedmann model. Communications in Mathematical Physics 46, 263–268.

Butterfield, J., 2006. Against *Pointillisme* about mechanics. British Journal for the Philosophy of Science 57, 709–753.

Butterfield, J., Isham, C., 2001. Spacetime and the philosophical challenge of quantum gravity. In: Callender, C., Huggett, N. (Eds.), Physics Meets Philosophy at the Planck Scale. Cambridge University Press, Cambridge, pp. 33–89.

Curiel, E., 1999. The analysis of singular spacetimes. Philosophy of Science 66 (Proceedings), S119–S145.

Cleland, C., 1984. Space: An abstract system of non-supervenient relations. Philosophical Studies 46, 19–40.

[59] What are these abstract elements, the relata of the relations, constituents of the structure? They are 'generalized' entities that are completely characterized in terms of (abstract) relational properties and that may correspond in some cases (that is, in mathematical terms, under a certain Gelfand (sub)representation) to the smooth real functions (possible scalar fields) and the usual tensor fields of the standard geometric formulation of GR. There is a deep issue here: fundamental physics makes reference to highly abstract mathematical structures and to some extent this fact may be used to vindicate (a version of) structuralism. But in order to give some content to such a structuralist position, it seems to be required to say something about the link between mathematical structures and physical ones.

[60] See Butterfield and Isham (2001, §2.2.2).

[61] There may be some analogy with Redhead's notion of 'surplus structure'.

[62] See also the conclusion in Bain (2006).

Demaret, J., Heller, M., Lambert, D., 1997. Local and global properties of the world. Foundations of Science 2, 137–176.
Dorato, M., 1998. Review of (Earman, 1995). British Journal for the Philosophy of Science 49, 338–347.
Dorato, M., 2000. Substantivalism, relationism, and structural spacetime realism. Foundations of Physics 30, 1605–1628.
Earman, J., 1987. Locality, nonlocality and action at a distance: A skeptical review of some philosophical dogmas. In: Kargon, R., Achinstein, P. (Eds.), Kelvin's Baltimore Lectures and Modern Theoretical Physics. Massachusetts Institute of Technology Press, Cambridge.
Earman, J., 1995. Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes. Oxford University Press, Oxford.
Earman, J., 1996. Tolerance for spacetime singularities. Foundations of Physics 26, 623–640.
Esfeld, M., Lam, V., 2008. Moderate structural realism about space-time. Synthese 160, 27–46.
French, S., Ladyman, J., 2003. Remodelling structural realism: Quantum physics and the metaphysics of structure. Synthese 136, 31–56.
Geroch, R., 1972. Einstein algebras. Communications of Mathematical Physics 26, 271–275.
Geroch, R., Horowitz, G., 1979. Global structure of spacetime. In: Hawking, S., Israel, W. (Eds.), General Relativity. Cambridge University Press, Cambridge.
Geroch, R., Liang, C., Wald, R., 1982. Singular boundaries of space-times. Journal of Mathematical Physics 23, 432–435.
Hawking, S., Ellis, G., 1973. The Large Scale Structure of Space-Time. Cambridge University Press, Cambridge.
Heller, M., 1992. Einstein algebras and general relativity. International Journal of Theoretical Physics 31, 277–288.
Heller, M., 2001. The classical singularity problem—history and current research. In: Martinez, V., Trimble, V., Pons-Borderia, M. (Eds.), Historical Development of Modern Cosmology. In: ASP Conferences Series, vol. 252. Astronomical Society of the Pacific, San Francisco, pp. 121–145.
Heller, M., Sasin, W., 1999. Origin of classical singularities. General Relativity and Gravitation 31, 555–570.
Johnson, R., 1977. The bundle boundary in some special cases. Journal of Mathematical Physics 18, 898–902.
Lam, V., 2008. The singular nature of spacetime. Philosophy of Science (PSA 2006 Proceedings), paper accepted.
Langton, R., Lewis, D., 1998. Defining 'intrinsic'. Philosophy and Phenomenological Research 58, 333–345.
Lewis, D., 1986. Philosophical Papers. Vol. 2. Oxford University Press, Oxford.
Mattingly, J., 2001. Singularities and scalar fields: Matter theory and general relativity. Philosophy of Science 68 (Proceedings), S395–S406.
Rickles, D., French, S., Saatsi, J. (Eds.), 2006. The Structural Foundations of Quantum Gravity. Oxford University Press, Oxford.
Schmidt, B., 1971. A new definition of singular points in general relativity. General Relativity and Gravitation 1, 269–280.
Rovelli, C., 2004. Quantum Gravity. Cambridge University Press, Cambridge.
Scott, S., Szekeres, P., 1994. The abstract boundary—a new approach to singularities of manifolds. Journal of Geometry and Physics 13, 223–253.
Senovilla, J., 1997. Singularity theorems and their consequences. General Relativity and Gravitation 29, 701–848.
Wald, R., 1984. General Relativity. The University of Chicago Press, Chicago.

CHAPTER 7

Who's Afraid of Background Independence?

Dean Rickles*

Abstract

Background independence is generally considered to be 'the mark of distinction' of general relativity. However, there is still confusion over exactly what background independence is and how, if at all, it serves to distinguish general relativity from other theories. There is also some confusion over the philosophical implications of background independence, stemming in part from the definitional problems. In this chapter I attempt to make some headway on both issues. In each case I argue that a proper account of observables goes a long way in clarifying matters. Further, I argue, against common claims to the contrary, that the fact that these observables are relational has no bearing on the debate between substantivalists and relationalists, though I do think it *recommends* a structuralist ontology, to as I endeavour to explain.

1. INTRODUCTION

> Everybody says they want background independence, and then when they see it they are scared to death by how strange it is … Background independence is a big conceptual jump. You cannot get it for cheap … (Rovelli, 2003, p. 1521)

In his 'Who's Afraid of Absolute Space?', John Earman (1970) defended Newton's postulation of absolute, substantival space at a time when it was very unfashionable to do so, relationalism being all the rage. Later, in his *World Enough and Space-Time*, he argued for a *tertium quid*, fitting neither substantivalism nor relationalism, substantivalism succumbing to the hole argument and relationalism offering "more promissory notes than completed theories" (Earman, 1989, p. 195). More recently, in a pair of papers written with Gordon Belot (1999, 2001), sub-

* History and Philosophy of Science, The University of Sydney, Australia

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00007-7
© Elsevier BV
All rights reserved

stantivalism comes under attack again. This time the culprit is the background independence of general relativity, and the potential background independence of a future theory of quantum gravity. The claim is that were the successful future theory of quantum gravity shown to be background independent, then substantivalism would be rendered untenable *for reasons of physics*—thus providing a clear-cut example of Shimonyan 'experimental metaphysics' in action.[1] Still more recently, in Volume 1 from this series (Dieks, 2006), Earman returns to his *tertium quid* idea, defending, again on the basis of (a manifestation of) background independence, what I consider to be a *structuralist* position which denies the fundamental existence of *subjects* (in the sense of 'bearers' of properties), thus ruling out both relationalism *and* substantivalism (Earman, 2006a).[2]

However, my primary target in this paper is the issue of what background independence is: only when this is resolved can we assess the claim that it might serve to settle debates over the ontology of spacetime (my secondary target). Let us begin by considering some basic metaphysical aspects of background structure, dependence, and independence, before firming the discussion up with the technical and definitional aspects. Ontological implications must wait until the final section.

2. METAPHYSICS OF BACKGROUNDS

Metaphysicians like to tell the following story to distinguish between physicalism and other non-physicalist positions:

> When God made the world did He lay out all the local physical matters of fact (properties at spacetime points) and the rest (causation, laws, modality, consciousness, etc.) followed, or did He then have to add these *after* or *in addition* to doing that?[3]

Physicalists, of course, think that in fixing the local physical facts in a world He thereby fixed *everything* there is to that world: all that there is physical. *Mutatis mutandis*, we can use this strategy to distinguish between positions on spacetime ontology too:

> When God made the world did He first create spacetime *and then* add matter (particles, fields, strings, etc.) to it or did He create matter and *thereby* fix the existence of spacetime?

[1] In fact, this is really just the hole problem again. In another paper, (Rickles, 2005b), I explicitly translated the hole argument into the framework of (background independent) loop quantum gravity, thus demonstrating that (this approach to) quantum gravity does not put the debate between substantivalists and relationalists on better ground than in the classical theory: substantivalists have nothing to fear from quantum gravity (not in the case of loop quantum gravity at any rate). I will aim to strengthen this conclusion further in this chapter.

[2] Earman goes so far as to suggest an entirely new ontological category: a "*coincidence occurrence*" (Earman, 2006a, p. 16). This is close to what I take to be one of the main ontological implications of background independence; however, as I have already intimated, I couch matters in structuralist terms—see Section 4. See also Rickles (2006) for a similar proposal drawn from the frozen formalism and problem of time in classical and quantum gravity, and Rickles (2008) for a more general defense of the view on the basis of (gauge) symmetries in physics.

[3] In other words, are causation, laws, modality, consciousness, etc., *supervenient* on the local physical matters of fact (but not *vice versa*), or do they constitute something 'over and above' these facts?

Or, in other words, do spacetime, and spatiotemporal properties and relations, exist *independently* of physical, material objects (particles, fields, strings, branes, etc...) or is the existence of some such objects *necessary*[4] for their existence? Substantivalists will answer Yes to the first disjunct and relationalists will answer Yes to the second. Let us be clearer on exactly what is meant by these terms. Here I shall follow Sklar (1974) (himself followed by Earman and a generation of philosophers of physics) in taking substantivalism to be the position that views spacetime to be an entity which exists *over and above* any material objects it might contain; or, in Earman's words, "prior to the objects it contains" instead of being "nothing but (might be constituted by, might be reducible to) the mutual relations among coexistent objects" (Earman, 1989, p. 289).

This also captures much of the intuitive distinction between background independence and dependence: is spacetime (geometry) fixed 'prior' to the determination of the state of matter in the universe or does one need to know what the state of matter is 'prior' to the determination of spacetime geometry? Given this superficial similarity, the distinction between background independence and background dependence is often supposed to latch on to the distinction between relationalism and substantivalism: relationalism being committed to the former; substantivalism being committed to the latter.[5] However, the 'dual role' of the metric field in general relativity rather muddies the waters here.

The schizophrenic nature of the metric field was viewed by Einstein as a necessary consequence of the equivalence principle, identifying inertia and weight: "the symmetric 'fundamental tensor' [$g_{\mu\nu}$] determines the metrical properties of space, the inertial behaviour of bodies in it, as well as gravitational effects" (Einstein, 1918c, p. 241). Or, as Carlo Rovelli puts it: "What Einstein has discovered is that Newton had mistaken a physical field for a background entity. The two entities hypostatized by Newton, space and time, are just a particular local configuration of a physical entity—the gravitational field—very similar to the electric and the magnetic field" (Rovelli, 2006, p. 27). In other words: "Newtonian space and time and the gravitational field are the same entity" (ibid.).

This duality, one expression of background independence in general relativity, has been responsible for much recent debate in the philosophy of spacetime physics. Again, it is seen to be implicated in the traditional debate between substantivalism and relationalism:

> In Newtonian physics, if we take away the dynamical entities, what remains is space and time. In relativistic physics, if we take away the dynamical entities, nothing remains. As Whitehead put it, we cannot say that we can have spacetime without dynamical entities, anymore than saying that we can have the cat's grin without the cat. (Rovelli, 2006, p. 28)

The reason being that the gravitational field is dynamical and does the work of two in also supplying the structures that characterize spacetime. However, the proposed link to the substantivalism-relationalism debate is problematic. Rovelli

[4] Or just *sufficient* if we wish to wage a war over ontological parsimony.

[5] There are other, *prima facie* more substantive reasons for the alignment; however, these reasons ultimately fail as well: see Rickles (2005b, 2005a, 2006, 2008) for the reasons why.

is lumping all of the dynamical fields together, as being ontologically 'on all fours'; but this is a mistake: we can remove all fields with the exception of the gravitational field and still have a dynamically possible world[6]—i.e. there are *vacuum* solutions to Einstein's field equations. But we cannot remove the gravitational field in the same way, leaving the other fields intact. This is an indication that there is something *special* about the gravitational field: it can't be switched off; it is not just one field among many.

Hence, the substantivalist would be perfectly within her rights to claim ownership. But, so would the relationalist since there is this ambiguity over the ontological nature of the field[7]: spacetime or material object? I take this state of affairs (namely 'joint ownership' of the metric field) to lend substantial support to Robert Rynasiewicz's (1996) claim that the debate between substantivalism and relationalism is "outmoded" in this context. However, we are drifting somewhat from our brief, which is to get a grip on the concept of background independence (and its companions, background structure and background dependence).

Background structures are contrasted with *dynamical* ones, and a background independent theory only possesses the latter type—obviously, background dependent theories are those possessing the former type in addition to the latter type.[8] Philosophers, and some physicists, will be more familiar with the term 'absolute element' in place of background structure, and the latter concept certainly soaks up a large part of the former. On the former concept, in his 1921 Princeton Lectures on the Theory of Relativity, Einstein writes:

> Just as it was consistent from the Newtonian standpoint to make both the statements, *tempus est absolutum, spatium est absolutum*, so from the standpoint of the special theory of relativity we must say, *continuum spatii et temporis est absolutum*. In this latter statement *absolutum* means not only "physically real," but also "independent in its physical properties, having a physical effect, but not itself influenced by physical conditions." (Einstein, 1921, p. 315)

There are three components here: a realist thesis, an independence thesis, and a non-dynamical thesis. Clearly the realist thesis can't simply mean that space and time exist, for Leibniz too would surely assent to such a thesis in *some* sense. Instead, I take it to mean that space and time are *fundamental* in the sense that they do not supervene on any further, underlying objects, properties, or facts. The independence thesis just looks like a denial of relationism, while the non-dynamical

[6] The still Machian Einstein of 1918 would not agree with this claim. He writes that "with [Mach's Principle] according to the field equations of gravitation, there can be no G-field without matter" (Einstein, 1918b, p. 34)—of course, this is where his λ-term appears, precisely in order to make the field equations compatible with Mach's Principle. The field equations lose out, being transformed into $G_{\mu\nu} - \lambda g_{\mu\nu} = -\kappa(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T)$, which do not allow for empty (i.e. $T_{\mu\nu} = 0$) spacetimes. The position I come to defend is not a million miles away from this: other fields are needed to form the gauge-invariant correlations (between field values) that provide the basic physical content of the theory.

[7] For example, the relationalist might, as Rovelli does, draw attention to the fact that "a strong burst of gravitational waves could come from the sky and knock down the rock of Gibraltar, precisely as a strong burst of electromagnetic radiation could" (Rovelli, 1997, p. 193).

[8] There is often a fair amount of slipping and sliding on this: there are *degrees* of background structure. Generally, one has in mind background *fields* rather than structures *per se*; that is, one is interested in the freedom (or not) from geometric-object fields on a manifold that are deemed 'background'. Though the manifold itself appears as a background structure, this is generally not counted when assessing a theory's background independence. This is a contentious point that we return to later—see, especially, footnote 10.

thesis amounts to 'absolute' in something like the sense of Anderson's notion of 'absolute object' (Anderson, 1967, pp. 83–87)—this itself, of course, corresponds most closely to Newton's notion of absolute in the sense of *immutability*, itself followed very closely by Einstein himself:

> If Newton called the space of physics 'absolute', he was thinking of yet another property of that which we call 'ether'. Each physical object influences and in general is influenced in turns by others. The latter, however, is not true of the ether of Newtonian mechanics. The inertia-producing property of this ether, in accordance with classical mechanics, is precisely *not* to be influenced, either by the configuration of matter, or by anything else. For this reason, one may call it 'absolute'. (Einstein, 1999, p. 15)

Anderson prefers to call this "the principle of reciprocity":

> It is seen that the absolute elements of a theory effect [*sic*.] the physical behaviour of a system. That is, a different assignment of values to the absolute elements would change the physical behaviour of the system. For instance, the assignment of different values to the metric might result in particle paths that are circles rather than straight lines. On the other hand, the physical behaviour of a system does not affect the absolute elements. An absolute element in a theory indicates a lack of reciprocity; it can influence the physical behaviour of the system but cannot, in turn, be influenced by this behaviour. This lack of reciprocity seems to be fundamentally unreasonable and unsatisfactory. We may express the converse in what might be called a general principle of reciprocity: Each element of a physical theory is influenced by every other element. In accordance with this principle, a satisfactory theory should have no absolute elements. (Anderson, 1964, p. 192)

Lee Smolin too adopts a similar line, explicitly linking this notion of absolute with the notion of background. He writes that "[t]he background consists of presumed entities that do not change in time, but which are necessary for the definition of the kinematical quantities and dynamical laws" (Smolin, 2006, p. 204). However, matters are not so simple as this. This rough 'absolute elements' way of defining background independence and background dependence is too vague to do any real work, and the various methods of firming things up face serious problems (as we shall see). Moreover, the peculiar nature of general relativity, replete with its treatment of the metric as a local dynamical variable, threatens to collapse the debate between substantivalism and relationalism. The interpretation of spacetime physics appears to be floundering.

Yet both debates, between background independence and background dependence and between substantivalism and relationalism, are believed by many physicists and a handful of philosophers to play a vital role in the search for a quantum theory of gravity. For example, in much of his recent work Lee Smolin (e.g. Smolin (2004, 2006)) defends the idea that background independence is a *necessary* piece of the quantum gravity puzzle: it is essential to solve the puzzles that quantum gravity raises that the geometry of spacetime is given as a solution of

some equations of motion, rather than placed in the theory 'by hand.' But Smolin also argues that background independence *uniquely* supports relationalism, claiming that physicists "often take background independent and relational as synonymous" (Smolin, 2006, p. 204). A big target in this paper is just this claim—a claim also made by Belot and Earman (1999, 2001) in order to prop up the listless body of the substantivalism/relationalism debate. *Substantivalists needn't be afraid of background independence any more than relationalists*. However, ultimately both lose out to a structuralist position!

3. DEFINITIONS AND DISPUTATIONS

It is often claimed that the *novelty* of general relativity lies in its (manifest) 'background independence.' However, background independence is a slippery concept apparently meaning different things to different people. In this section we attempt to gain a firmer grip on this slippery customer by considering various elucidations of background independence that have been suggested. There is a clear core to the notion, and I argue that this core can be made clearer by connecting the concept of background independence to the nature of the observables in background independent theories.

Let us begin by presenting a general way of making sense of the various proposals—here I largely follow (Giulini, 2006). Let us specify a theory by writing down its laws as a set of equations of motion representing relations between the central objects of the theory. We get the following schema:

$$\mathfrak{E}[\mathcal{D}, \mathcal{B}] = 0 \tag{1}$$

Here $\mathcal{D}$ represents the dynamical structures (those that have to be solved to get their values, such as the electromagnetic field and the metric in general relativity—these represent the physical degrees of freedom of the theory, out which the observables will be constructed) and $\mathcal{B}$ the background structures (those whose values are put in by hand, such as the topology and, in pre-general relativistic theories, the metric). Now let us represent the space of *kinematically* possible histories by $\mathcal{K}$. Then $\mathfrak{E}[\mathcal{D}, \mathcal{B}] = 0$ selects a subset $\mathcal{P} \subset \mathcal{K}$ of *dynamically* possible histories (or 'physically' possible worlds) relative to $\mathcal{B}$.[9] Now, if there are no such $\mathcal{B}$s (or, rather, no $\mathcal{B}$*-fields*) then the physically possible histories (the dynamics) is given by relations between the $\mathcal{D}$s (and, at least fiducially—i.e. in terms of the *formal* definition of the fields—the manifold, but the diffeomorphism symmetry washes this dependence away). This impacts on the observables of the theory, for the observables must then make no reference to the $\mathcal{B}$s, only to the $\mathcal{D}$s. This is the source of the claim that general relativity, and background independent theories, are *relational*: it simply means that the states and observables of the theory do not make reference to background structures.[10]

[9] As Wheeler puts it, "[k]inematics describes conceivable motions without asking whether they are allowed or forbidden. Dynamics analyses the difference between a physically reasonable and a disallowed history" (Wheeler, 1964, p. 65).

[10] Though, again, this does not include the manifold which is required for the (formal) definition of the dynamical fields. The inescapable presence of the manifold, in which dimension, topology, differential structure and signature are fixed in-

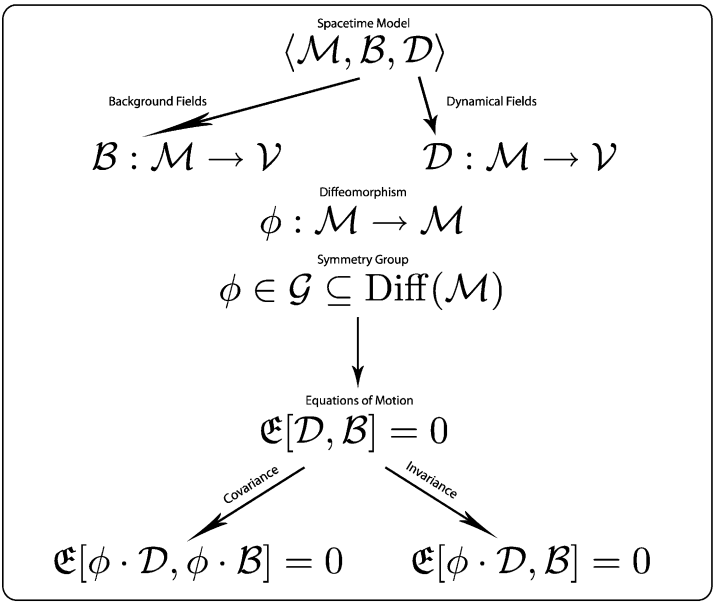

FIGURE 7.1 How to understand covariance and a invariance groups in a spacetime theory. Here, the fields take values in a vector space (or a more structured space). The diffeomorphisms drag fields along to new points. The equations of motion are of the form 'solve for $\mathcal{D}$ given $\mathcal{B}$'.

This way of understanding a theory lets us recapitulate in a clearer way our earlier definitions of covariance and invariance. Let $\mathcal{G}$ be a group of spacetime symmetries that acts on $\mathcal{K}$ as $\mathcal{G} \times \mathcal{K} \to \mathcal{K}$—i.e. elements of $\mathcal{G}$ map kinematically possible solutions onto kinematically possible solutions. We say that $\mathcal{G}$ is a symmetry group of the theory whose space of kinematically possible histories is $\mathcal{K}$ just in case $\mathcal{P}$ is left invariant by its action. Alternatively, and more usefully for what follows, we can express the distinction between covariance and invariance as—this is summarized schematically in Figure 7.1.[11]

$$[\mathbf{COV}] \Rightarrow \mathfrak{E}[\mathcal{D}, \mathcal{B}] = 0 \quad \text{iff} \quad \mathfrak{E}[g \cdot \mathcal{D}, g \cdot \mathcal{B}] = 0 \quad (\forall g \in \mathcal{G}) \tag{2}$$

$$[\mathbf{INV}] \Rightarrow \mathfrak{E}[\mathcal{D}, \mathcal{B}] = 0 \quad \text{iff} \quad \mathfrak{E}[g \cdot \mathcal{D}, \mathcal{B}] = 0 \quad (\forall g \in \mathcal{G}) \tag{3}$$

dependently of the equations of motion, leads Smolin to call general relativity only a "partly relational theory" (Smolin, 2006, §7.4). However, the absence of background fields coupled with the symmetry of the manifold means that a displacement (*via* a diffeomorphism) of the dynamical fields with respect to it simply produces a gauge-equivalent representation of one and the same physical state. Elimination of these redundant possibilities ("surplus structure" in Redhead's sense (1975)) further reduces the size of $\mathcal{P}$, giving us the reduced space $\overline{\mathcal{P}} = \mathcal{P}/\mathrm{Diff}(\mathcal{M})$. This 'superspace' contains points that are entire orbits of the gauge group, representing *abstract* 'delocalized' structures known as a "geometries"—see Misner et al. (1973, p. 522). This is supposed to be a space fit only for relationalists; however, there are plenty of good arguments that show that the substantivalist has just as much right to occupy it—see Pooley (forthcoming) and Rickles (2008) for more details.

[11] Cf. Giulini (2006, p. 6). I recommend that all philosophers of physics interested in background independence, and the difficulties in defining absolute objects, read this article: it is an exceptionally clear-headed review.

As I said, in the context of general relativity **[COV]** = **[INV]** since $\mathcal{B} = \emptyset$ (manifold aside). Of course, the fact that the manifold appears in the laws—and the absence of symmetry-reducing background fields (i.e. to reduce the effective symmetry group of the theory)—means that the there will be surplus structure: the localization on the manifold of the dynamical fields is pure gauge.[12]

All this symmetry affects the dynamics so that a standard Hamiltonian or Lagrangian formulation is not possible. Respectively, the canonical variables are not all independent (being required to satisfy identities known as constraints: $\phi(q,p) = 0$) and the Euler–Lagrange equations are not all independent. These identities serve to 'constrain' the set of phase space points that represent genuine physical possibilities: only those points satisfying the constraints do so, and these form a subset in the full phase space known as the 'constraint surface'.

As I also said, this has an impact on the form of the observables—and this is terribly important for the quantization of the theory. Since a pair of dynamical variables (not observables) that differ by a gauge transformation are indistinguishable, corresponding to one and the same physical state of affairs, the observables ought to register this fact too: that is, the observables of a gauge theory should be *insensitive* to differences amounting to a gauge transformation—as should the states in any quantization of such a theory: i.e. if $x \sim y$ then $\Psi(x) = \Psi(y)$.[13] Where '$\mathcal{O}$' is a dynamical variable, 'O' is the set of (genuine) observables, $x, y \in \mathcal{P}$, and '$\sim$' denotes gauge equivalence, we can express this as:

$$\mathcal{O} \in \mathrm{O} \iff (x \sim y) \supset (\mathcal{O}(x) = \mathcal{O}(y)) \tag{4}$$

Or, equivalently, we can say that the genuine observables are those dynamical variables that are constant on gauge orbits '$[x]$' (where $[x] = \{x\colon\ x \sim y\}$):

$$\forall [x],\ \mathcal{O} \in \mathrm{O} \iff \mathcal{O}[x] = \text{const.} \tag{5}$$

Most of the work done on finding the observables of general relativity is done using the $3+1$ projection of the spacetime Einstein equations. That is, the constraints are understood as conditions laid down on the initial data[14] $\langle \Sigma, h, K \rangle$ when we project the spacetime solution onto a spacelike hypersurface Σ—here, h is a Riemannian metric on Σ and K is the extrinsic curvature on Σ. I won't go into the nitty gritty details here, but it turns out that the Hamiltonian of general relativity is a sum of constraints on this initial data (of the kind that generate gauge motions,

[12] This difference corresponds, then, to that between 'passive' and 'active' diffeomorphism invariance. As Rovelli puts it: "A field theory is formulated in manner invariant under passive diffs (or change of co-ordinates), if we can change the coordinates of the manifold, re-express all the geometric quantities (dynamical *and non-dynamical*) in the new co-ordinates, and the form of the equations of motion does not change. A theory is invariant under active diffs, when a smooth displacement of the dynamical fields (*the dynamical fields alone*) over the manifold, sends solutions of the equations of motion into solutions of the equations of motion" (Rovelli, 2001, p. 122). We will call the former general covariance and the latter diffeomorphism invariance—Earman (2006b) calls the latter *substantive* general covariance, on the understanding that it amounts to a gauge symmetry, as we have assumed.

[13] It seems that Einstein was aware of this implication soon after completing his theory of general relativity, for he writes that "the connection between *quantities in equations* and *measurable quantities* is far more indirect than in the customary theories of old" (Einstein, 1918a, p. 71).

[14] Note that John Wheeler refers to constraints as "initial value equations" (Wheeler, 1964, p. 83). This terminology gets one closer to the *physical* meaning of the constraints.

namely 1st class)—hence, the dynamics is entirely generated by constraints and is therefore pure gauge.[15]

This formulation allows us to connect the characterization of the observables up to the dynamics (generated by constraints $\mathcal{H}_i$) more explicitly:

$$\mathcal{O} \in \mathrm{O} \iff \{\mathrm{O}, \mathcal{H}_i\} \approx 0 \; \forall i \tag{6}$$

In other words, the observables of the theory are those functions that have weakly vanishing (i.e. on the constraint surface) Poisson brackets with all of the first-class constraints. These are the gauge-invariant quantities. A pressing problem in general relativity—especially pressing for quantum gravity—is to find suitable entities that satisfy this definition. There are at least two types that fit the bill: highly non-local quantities defined over the whole spacetime[16] and (differently) non-local, 'relational' quantities built out of correlations between field values. There seems to be some consensus forming, at least amongst 'canonical relativists', that the latter type are the most natural.

I return to the interpretation of these correlational observables in Section 4. Let us now consider a number of standard takes on the question of what background independence amounts to. We assess two main ways of doing this—utilizing general covariance and diffeomorphism invariance—before considering J.L. Anderson's proposal for firming up the latter approach and discussing the connection to observables.

As we saw above, general covariance simply refers to the fact that when we hit a solution with an arbitrary diffeomorphism, we get another solution back. That is to say, the equations of motion are covariant with respect to diffeomorphisms. This amounts to a carrying along of all of the fields. Covariance is not so restrictive as invariance (or "symmetry" as Anderson calls it). The former just says that if M is a solution then so is M where $\mathrm{M}' = g(\mathrm{M})$ and g is an element of the covariance group, the group that preserves the form of the laws (the equations of motion). The theory is said to be g-covariant. But this is just a constraint on the *form* of the theory, not on its physical content. In other words, general covariance in this sense is simply a property of the formulation of the theory. This is, of course, just what Kretschmann taught Einstein soon after general relativity was written down in its final form. The problem is that (general) covariance just means that the equation of motion is *well-defined* in the sense of differential geometry: the equation need simply 'live' on the manifold.

Recall how coordinatization occurs. Firstly, we associate points of the manifold with $\mathbb{R}^4$ so that each $x \in \mathcal{M}$ gets four numbers $\{x^\mu\}$ ($\mu = 1, 2, 3, 4$) associated to it. We can do this in many ways, as mentioned above. We might use the assignment $\{x'^\mu\} \to y \in \mathcal{M}$ instead. Because these numbers are assigned to the *same* point there

[15] This turns out to be behind the two worst conceptual problems of general relativity: the hole argument and the frozen formalism problem. For details on these connections see Belot and Earman (1999, 2001) and Rickles (2006). Earman (2003) gives a splendid presentation of the relationship between the constrained Hamiltonian formalism and gauge, including their implications for time and change.

[16] There is a proof (for the case of closed vacuum solutions of general relativity) that there can be no *local* observables at all (Torre, 1993)—'local' here means that the observable is constructed as a spatial integral of local functions of the initial data and their derivatives.

will be some relationship between the coordinate systems:[17]

$$x'^{\mu} = x'^{\mu}(x^{\nu}) \tag{7}$$

Given the differential structure of the manifold, we get an infinitely continuously differentiable function between the thus related coordinate systems (with a similarly differentiable inverse).[18] This is a diffeomorphism passively construed; it is gauge in a very trivial sense, as Wheeler says: "How one draws coordinate surfaces through space-time is a matter of paperwork and bookkeeping, and has nothing to do with the real physics" (Wheeler, 1964, p. 81). This goes for any reasonable spacetime theory. Hence, general covariance understood in these terms does not have the power to distinguish between spacetime theories, and if background independence is supposed to distinguish general relativity from previous theories, then general covariance cannot underwrite it.

Hence, *any* spacetime theory written in terms of geometric objects on the manifold will be generally covariant in the sense of having Diff($\mathcal{M}$) as its covariance group. However, *invariance* is a much stronger requirement that picks out a subgroup of the covariance group; this says that if M is a solution then so is M where $\mathrm{M}' = g(\mathrm{M})$ and now g is an element of a subgroup of the covariance group that preserves the absolute elements. The theory is said to be g-invariant. Hence, in one we map all of the objects, in the other we only map the dynamical objects, whilst preserving the background structure.[19]

This is the standard view: rather than considering the background fields to be transformed along with the dynamical fields, we view the diffeomorphisms in an *active* way, as shifting the dynamical fields relative to the *same* background fields. Thus, Lee Smolin writes that "[g]eneral coordinate invariance is not the same thing as diffeomorphism invariance, and it is the latter, and not the former, that is the key to the physical interpretation of the theory". He goes on to say that

> with the introduction of explicit background fields any field theory can be written in a way that is generally coordinate invariant. This is not true of diffeomorphism invariance, which relies on the fact that in general relativity there are no non-dynamical background fields. Diffeomorphisms, in contrast to general coordinate transformations, are active transformations that take points to other points, so that diffeomorphism invariance is, explicitly, the statement that the points are not meaningful. Both philosophically and mathematically, it is diffeomorphism invariance that distinguishes general relativity from other field theories. (Smolin, 2003, p. 234)

[17] It is useful to think of the pair of coordinate systems as being like a pair of languages and as the particular coordinates assigned to some particular point as being like nouns in the language referring to a particular object. A translation manual between the languages would be analogous to the differentiable functions relating distinct coordinate descriptions of one and the same point; in this case we would have distinct words referring to the same object, these words being inter-translatable.

[18] Generally, because the manifolds in general relativity are coordinatized by gluing patches together, the function will be evaluated on the *overlap* between coordinate systems.

[19] The story then goes: special relativity cannot be diffeomorphism invariant—i.e. it cannot have Diff($\mathcal{M}$) as its symmetry group—because the imposition of the Minkowski metric reduces the invariance group to a subgroup of the covariance group, namely the Poincaré group of isometries of this metric, of Minkowski spacetime. This smaller group is the largest that preserves the (background) structure of Minkowski spacetime; there are clearly elements of Diff($\mathcal{M}$) that would not do so.

There are at least two ways in which this misses the mark. Firstly, one can retain the physical content of diffeomorphism invariance without disposing of points: either one can adopt a Kretschmann–Bergmann–Komar 'intrinsic coordinates' method ((Kretschmann, 1917; Bergmann, 1977, 1961; Bergmann and Komar, 1972)—see also Section 4), or else one can view diffeomorphism invariance as imposing a constraint on the form of the observables of a theory, so that it is true that "points are not meaningful" only in the sense that *from the point of view of the physics*, as encoded in the observable content of the theory, there is an 'indifference' to the points of the manifold (i.e. as to which point plays which role). Both are compatible with there being points. Secondly, and more problematic for our purposes, is that this faces a Kretschmann-type objection too: *any* background field can be made dynamical by making it satisfy some equations of motion, however physically vacuous they might happen to be. Hence, unless we have some other way of making sense of the distinction between background and dynamical fields, then this account fails in the same way the general covariance account fails—fortunately, Anderson provides just such a method, but first let us go through the details of why this account fails as it stands.

The objection is that we are free to extend the invariance group to the covariance group by making any background fields into dynamical ones, thus collapsing the distinction between invariance and covariance groups. If there are no background fields then the invariance group automatically becomes identical to the covariance group (i.e. the diffeomorphism group). In the case of a specially relativistic theory, say, in order to preserve the structure of Minkowski spacetime we would have to impose a condition of flatness on the metric. But, of course, this makes the metric dynamical (in the sense of satisfying equations of motion)! The problem is, this all depends upon the availability of some way of distinguishing between absolute and dynamical fields, and so far we simply have an intuitive notion. Clearly if this intuitive notion amounts to 'being solved for' then we can make special relativity background independent, which then conflicts with our basic intuitions about what background independence is.

Take the following stock example of a massless scalar field on Minkowski spacetime:

$$\Box\phi \equiv \eta^{\mu\nu}\nabla_\mu\nabla_\nu\phi = 0 \tag{8}$$

All we do here is replace the background metric with a general metric and make the new metric obey a 'flatness condition':

$$\eta_{\mu\nu} \to g_{\mu\nu} \tag{9}$$

$$\Box\phi \equiv g^{\mu\nu}\nabla_\mu\nabla_\nu\phi = 0 \tag{10}$$

$$\mathrm{Riem}[g] = 0 \tag{11}$$

Like the generally relativistic case, we now appear to have no background fields! If diffeomorphism invariance is what underwrites background independence then

the latter cannot be what makes general relativity *special*.[20] Hence, it appears that we have made a specially relativistic theory generally relativistic!

Anderson (1964, pp. 182–3) complains about this procedure on the grounds that the way the general metric was introduced was physically unmotivated: there is no need to have a general metric since nonflat metrics are not considered. Presumably this is similar to what Einstein had in mind when he complained that although it was possible to reformulate other spacetime theories in a generally covariant way, this does not produce their *simplest* formulation: "Among two theoretical systems, both compatible with experience, one will have to prefer the one that is simpler and more transparent from the point of view of the absolute differential calculus" (Einstein, 1918b, p. 34). There is a lot of room for the metric to move with a general metric that is not being occupied in these reformulations; hence, no additional content is being added by using it. Anderson argues that the expansion of the invariance group of a theory (to encompass a larger covariance group of which it is usually a subgroup) reveals pre-existing absolute objects in the theory. The metric in special relativity is an absolute object whose "existence was masked by the fact that [it] had been assigned [a] particular [value]" (Anderson, 1964, p. 192).

Anderson provides the split between background and dynamical fields so that the diffeomorphism invariance definition of background independence can do its work. What Anderson proposes is that absolute elements (called "absolute objects" in Anderson (1967)), understood as variables whose "determination is entirely independent of other physical objects of the theory" (Anderson, 1964, p. 186), serve to define "the relativity principle" associated with some theory—that is, the principle stating that the theory's laws are invariant under the invariance group. Let us work through Anderson's proposal to see how this works and how it contributes to the problem of defining background independence.

He begins, as we did above, by specifying a theory as a set of functional relations (i.e. equations of motion) between the independent variables of the theory (particles, fields, fluids, strings, branes, etc.):

$$\mathcal{L}_i(y_A) = 0 \tag{12}$$

Anderson seeks to find a way of identifying the background structures in a theory that is so specified. The idea is to inspect all of the invariant functions (the *genuine* observables) under the theory's covariance group that can be constructed from various subsets of the variables y_A, and to then see if the values of these functions are uniquely determined by equation (12) alone (independently of additional conditions). Those values of the functions that are determined independently of the values of others are deemed absolute.

There is something *prima facie* rather peculiar about Anderson's analysis in that it implies that the metric $g_{\mu\nu}$ of general relativity constitutes an absolute element (with the Lorentz group as its invariance group) in the vacuum case, because it is

[20] As Guilini notes (Giulini, 2006, pp. 13–14), there are in fact problems with this example: if we consider the (reasonable) requirement that our equations of motion have to be the Euler–Lagrange equations for some action principle then we find that the action principle delivering equations (10) and (11) generates a bigger solution space than that of equation (8). The two are *not* equivalent formulations of the same theory.

uniquely determined (up to diffeomorphism) *independently of other physical objects*, but not in the matter-present case, where it is determined by other physical objects in the theory. Hence, it isn't an absolute matter whether the metric in general relativity is an absolute object or not; rather, it depends on whether there are other fields present, and so on which field equations are appropriate. This is at odds with what we might expect, namely a definition of background independence that renders general relativity background independent *simpliciter*. However, given what I had to say about the observables—i.e. that they are correlations between fields values—I think this is what we *should* say.[21]

Background independence is, then, defined using this machinery: a theory is background independent just in case it contains no absolute elements. This lines up with the diffeomorphism invariance account, for a diffeomorphism invariant theory will have no background structures; this how we get the identity between the covariance and invariance groups. This clearly renders general relativity background independent; its covariance group is indeed identical to its invariance group (or its 'relativity group'). The method gets the relativity principles for other spacetime theories right too.

If we turn what were originally background fields into dynamical fields, by making them obey equations of motion (in the sense of Anderson), then they will enter into the definition of the observables, since we are understanding the $\mathcal{D}$ to be the 'ingredients' of the observables. This is how we end up with hole-type problems: background fields—however they are tweaked in an attempt to make them dynamical—introduce unobservable ('unphysical surplus') content into the physical structure (as given by the $\mathcal{D}$s). This can be seen explicitly if we consider how the born-again dynamical fields look in the solution space of the theory. $\mathcal{P}$ is our solution space, and since we are taking the theory to be diffeomorphism invariant it will carry an action of the gauge group Diff($\mathcal{M}$). The dynamical fields serve to separate the points of $\mathcal{P}$. However, there will be a redundancy in the labeling since the diffeomorphism invariance allows us to construct solutions from solutions by acting on them with elements of Diff($\mathcal{M}$). For each solution of the original equations we now have an orbit of solutions. If we understand the diffeomorphism symmetry as a gauge freedom then this will be a gauge orbit. This gives us a potential way to further detrivialize this approach, for we can see that the flatness condition forces the value of $g_{\mu\nu}$ to be the same in *each and every orbit*.[22] Hence, if we identify the gauge orbits then we have just one state here—I am shelving complications to do with locality; for more details see Giulini (2006, §2.5). This is what it means to say the metric is an absolute object: something that is the same in every solution. But the inability to distinguish between orbits is the definition of an unobservable here too, so we have the connection between background independence and observables that we were looking for (and the connection between

[21] Compare this with the Einstein quote I give on p. 149. I think this clearly shows that Einstein would have sided with Anderson on this point.

[22] That is, the value is not just constant on gauge orbits—which is part and parcel of being a good observable—it is constant *across* gauge orbits too. No observable can distinguish between such orbits; hence the structure is unobservable.

background structures and non-observability). Moreover, it matches what we intuitively mean by background independence.[23]

4. IMPLICATIONS FOR THE ONTOLOGY OF SPACETIME

Diffeomorphism invariance makes spacetime local observables an impossibility. Since there clearly are local degrees of freedom, and these are what we observe, we need some notion of local observable that does not make reference to spacetime geometry. That is, we need a background independent notion of local observable. The obvious (and indeed, only thing to do) is to use physical degrees of freedom to localize. The observables so localized are relational.

Calling those dynamical variables whose motion can be uniquely determined by the field equations the "true observables", Anderson writes that:

> A unique state of the system is … specified by giving, at some instant of time, values of the true observables and their first time derivatives. In a sense, these true observables are the physical meaningful "coordinates" of the system. (Anderson, 1958, p. 1197)

These true observables are the gauge-invariant quantities I mentioned earlier. Earman asks: "Does the gauge-invariant content of GTR characterize a reality that answers the relationist's dreams, or do the terms of the absolute-relational controversy no longer suffice to adequately describe what Einstein wrought?" (Earman, 2006a, p. 10). Earman answers Yes to the latter disjunct, and No to first. He proposes to put an entirely new ontological scheme, based on 'coincidence occurrences', in place of the absolute and relational positions. As Earman points out, "a coincidence occurrence consists in the corealization of values of pairs of (non-gauge invariant) dynamical quantities" (Earman, 2006a, p. 16). Earman thinks that this new conception of physical quantities signals the necessity of a shift from the traditional 'subject-predicate'-based ontologies, such as substantivalism and relationalism. As I said earlier, I think this is the right thing to say; however, I would spell it out rather differently, in terms of structuralism. Rovelli's framework of partial and complete observables provides the formal underpinning.

Firstly, how might relationalism and substantivalism get a foothold in this background independent context? According to the relationalist (about motion) all motion is relative motion. But motion relative to what? The gravitational field? But if it is the gravitational field, then we face a problem in GR (and background

[23] There are problem cases that remain, as discussed, for example, in Pitts (2006)—namely, the so-called 'Jones–Geroch counterexample' which apparently shows that the 4-velocity of a 'cosmic dust field' counts as a background structure (according to the Anderson–Friedman analysis). The problem stems from Michael Friedman's' modification of Anderson's identification of background structures as those with single $\mathrm{Diff}(\mathcal{M})$ orbits. Friedman argues that the condition should be made *local* in order to get at the notion that background structures do not correspond to local degrees of freedom. To achieve this he counts as background structures fields that are locally diffeomorphism equivalent—this condition is satisfied when there is a diffeomorphism mapping neighbourhoods (of any manifold point) to neighbourhoods, such that two fields restricted to the neighbourhoods (connected by a carry-along) take on the same values. The problem is that any pair of nowhere vanishing vector fields will always satisfy this condition and, therefore, always count as background structures. The absurd conclusion is that any diffeomorphism invariant vector field theory will automatically be branded background dependent. Utilizing the observables can help here, at least in the present case: the observables will register the physical fact that such fields will generally not cover the *whole* of spacetime.

independent theories in general): is this field spacetime or matter? Einstein, and Rovelli, claim that the gravitational field should be *identified* with spacetime. Here we see that both positions can get a foothold on the ontological rock face of general relativity; the substantivalist can lay claim to the same object against which relative motion occurs. The same goes for localization, which is, I suspect, more what Earman has in mind: if localization is relative to the gravitational field, then both substantivalists and relationalists (in the ontological sense) can get a foothold. Matters have clearly degenerated (pardon the pun) to the point where this division is no longer doing any real work.

But we can say more. I mentioned above that reduction (i.e. the elimination of symmetry) was supposed to be implicated. The idea here is that the natural representational tool for relational spacetime is the *geometry* rather than individual metrics on the manifold:

> [T]he basic postulate that makes GR a relational theory is [that] ... *[a] physical spacetime is defined to correspond, not to a single* (M, g_{ab}, f)*, but to an equivalence class of manifolds, metrics, and fields under the action of Diff(M).* (Smolin, 2006, p. 206)

The idea here, then, is that removing the symmetries (by 'quotienting out' by the diffeomorphisms) is taken to correspond to relationalism; or, in other words, that relationalism is reductionism. This is tantamount to the gauge-invariance view. It poses no obstacle for the substantivalist; there are a variety of ways to accept it, most of which amount to a denial of haecceitism of some sort or another (i.e. the claim that there can be worlds that differ non-qualitatively)—see Pooley (forthcoming) for more details. There is no necessity gluing haecceitism and substantivalism together, and a relationalist can just as well be an haecceitist.

Like Rovelli and Smolin, Wheeler dismisses the points of space as "[m]ere baggage". The coordinate representations we use hide the real, objective reality: the geometries. Hence, a geometry is an abstract object that encodes the intrinsic features of a space: it stands one-to-many with localized metrics. Again, there is no reason why the substantivalist shouldn't say that this intrinsic structure is what they mean by spacetime and where their ontological commitments lie. On the other side of the coin it is possible that the substantivalist can retain points in the face of the gauge freedom. As Robert Dicke remarks, speaking on behalf of J.L. Synge:

> general relativity describes an absolute space ... certain things are measurable about this space in an absolute way. There exist curvature invariants that characterize this space, and one can, in principle, measure these invariants. Bergmann has pointed out that the mapping of these invariants throughout space is, in a sense, labelling of the points of this space with invariant labels (independent of coordinate system). These are concepts of an absolute space, and we have here a return to the old notions of an absolute space. (Dicke, 1964, p. 124–125)

Here the idea is to get a set of coordinate conditions that allow one to define a set of intrinsic coordinates. One constructs the complete set of scalars from the metric

and its first and second derivatives, which for the matter-free case leaves four non-zero scalars that take different values at different points of the manifold. Hence, one achieves a complete labelling of the manifold in an intrinsic gauge-invariant—this follows from the fact that we are dealing with *scalars* which do not change their values under diffeomorphisms. These points can then be used to localize quantities which become gauge-invariant as a result of the gauge-invariance of the scalars. For Synge, the only difference between this space and Newton's is that the geometric properties of the Einsteinian space are "influenced by the matter contained therein"—that is, the latter is background independent. Of course, since we are dealing with invariants of the metric here, it is open to the relationalist to call this a material field. So continues the interminable tug-of-war!

I think this is evidence in favour of the view that the time has come to forget about the 'debate' between substantivalism and relationalism, and focus on an alternative. Here I argue that structuralism offers a suitable alternative. The position involves the idea that physical systems (which I take to be characterized by the values for their observables) are exhausted by extrinsic or relational properties: they have no intrinsic properties at all! This is a consequence of background independence coupled with gauge invariance. This leads to a rather odd picture in which objects and structure are deeply entangled in the sense that, inasmuch as there are objects, any properties they possess are structurally conferred: they have no reality outside the correlation. What this means is that the objects don't *ground* the structure; they are nothing independently of the structure, which takes the form of a (gauge-invariant) correlation between (gauge variant) field values.[24] We can sum this up by paraphrasing one of Hermann Minkowski's infamous remarks:

> *Henceforth spacetime by itself and matter by itself are doomed to fade away into mere shadows, and only some kind of union between the two can preserve their independent reality!*

This admittedly rather wild-sounding metaphysics can be made more precise through the use of Rovelli's framework of partial and complete observables.

A *partial* observable is a physical quantity to which we can associate a measurement leading to a number and a *complete* observable is defined as a quantity whose value (or probability distribution) can be predicted by the relevant theory. Partial observables are taken to coordinatize an extended configuration space $\mathcal{Q}$ and complete observables coordinatize an associated reduced phase space Γ_{red}. The "predictive content" of some dynamical theory is then given by the kernel of the map $f : \mathcal{Q} \times \Gamma_{red} \to \mathbb{R}^n$. This space gives the *kinematics* of a theory and the *dynamics* is given by the constraints, $\phi(q^a, p_a) = 0$, on the associated extended phase space $T^*\mathcal{Q}$. The content appears to be this: there are quantities that can be measured whose values are *not* predicted by the theory. Yet the theory *is* deterministic because it does predict correlations between partial observables. The dynamics is then spelt out in terms of *relations* between partial observables. Hence, the theory

[24] There is kinship here with Eddington who writes that "the significance of a part cannot be dissociated from the system of analysis to which it belongs. As a structural concept the part is a symbol having no properties except as a constituent of the group-structure of a set of parts" (Eddington, 1958, p. 145); and later, "a structure does not necessarily imply an X of which it is the structure" (Eddington, 1958, p. 151).

formulated in this way describes relative evolution of (gauge variant) variables as functions of each other. No variable is privileged as *the independent one* (cf. Montesinos et al., 1999, p. 5). The dynamics concerns the relations between elements of the space of partial observables, and though the individual elements do not have a well defined evolution, relations between them (i.e. correlations) do: they are independent of coordinate space and time.

Following Dittrich (2006), the interpretation here is as follows: $\phi = T$ is a partial observable parametrizing the ticks of a clock (laid out across a gauge orbit), and $f = a$ is another partial observable (also stretching out over a gauge orbit). Both are gauge variant quantities. A gauge *invariant* quantity, a complete observable, can be constructed from these partial observables as:

$$\mathcal{O}_{[f;T]}(\tau, x) = f(x') \tag{13}$$

These quantities encode correlations. They tell us what the value of a gauge variant function f is when, under the gauge flow generated by the constraint, the gauge variant function T takes on the value τ. This correlation is gauge invariant. These are the kinds of quantity that a background independent gauge theory like general relativity is all about. We don't talk about the value of the gravitational field *at a point of the manifold*, but where some other field (say, the electromagnetic field) takes on a certain value. Once again, we find that Einstein was surprisingly modern-sounding on this point, writing that "the gravitational field at a *certain location* represents nothing 'physically real,' but the gravitational field together with other data does" (Einstein, 1918a, p. 71).

Now here I would agree with Einstein and disagree with Rovelli about the interpretation of these correlations. Rovelli claims that "the *extended* configuration space has a direct *physical* interpretation, as the space of the partial observables" (Rovelli, 2002, p. 124013-1, my emphasis). *Both* spaces—the space of genuine (complete) observables and partial observables—are invested with physicality by Rovelli; the partial observables, in particular, are taken to be physical variables. Einstein argues that only the correlation is physically real in a fundamental sense. In this he is clearly followed by Stachel (1993) who argues that the kinematical state space of a background independent theory like general relativity has no physical meaning prior to a solution (so that only the dynamical state space is invested with the power to represent genuine physical possibilities; kinematics then being in this sense derivative).

It is for this reason that I think structuralism can help with the interpretation of background independent, gauge-invariant theories—that is, we don't need to go as far as Earman in postulating a whole new ontological category. Recall that epistemic structural realism argues that the best we can hope for is to get to know structural aspects of the world, since we only ever get to observe relational properties rather than intrinsic ones (in our experiments and so on). However, in a background independent gauge theory like general relativity we have seen that the physical observables just *are* relational quantities: this is all there is! In other words, there's nothing 'underneath' the relational properties (as encoded in the

$\mathcal{D}$-fields), so that these *exhaust* what there is, leading to an *ontological* structuralism.[25] This is why we face the problems regarding the 'subject-predicate'-style ontologies that Earman mentions: there *are* no independent subjects that are the 'bearers' of properties and the 'enterers' of relations. Hence, unless one can have objects without intrinsic properties (and I don't think this is a metaphysically healthy route to follow), we should follow Earman's lead, and I say that this journey will lead us to some variant of ontic structural realism.

5. CONCLUSION

I have argued that we can make good sense of background structure and background independence by following an Anderson-style account (involving the view that background structures have single Diff($\mathcal{M}$) orbits) and utilizing the appropriate gauge-theoretic definition of 'observable'. These set up a connection between Anderson's idea and the intuitive notion that background structures are not the kinds of thing we can measure, are not the kinds of entity that can ground those quantities we might wish to measure (field values and so on). The ontological implications of background independence, so conceived, are not what is often claimed: relationalism is not uniquely supported. Substantivalists too can uphold their interpretation in the context of background independent theories. However, aspects of the observable content of background independent theories were shown to cause problems for both relationalism and substantivalism. I argued that these aspects recommend a structuralist position.

REFERENCES

Anderson, J.L., 1958. Enumeration of the true observables in gauge-invariant theories. Physical Review 110 (5), 1197–1199.

Anderson, J.L., 1964. Relativity principles and the role of coordinates in physics. In: Chiu, H.-Y., Hoffmann, W.F. (Eds.), Gravitation and Relativity. W.A. Benjamin, Inc., pp. 175–194.

Anderson, J.L., 1967. Principles of Relativity Physics. Academic Press, New York.

Belot, G., Earman, J., 1999. From metaphysics to physics. In: Butterfield, J., Pagonis, C. (Eds.), From Physics to Metaphysics. Cambridge University Press, pp. 166–186.

Belot, G., Earman, J., 2001. Presocratic quantum gravity. In: Callender, C., Huggett, N. (Eds.), Physics Meets Philosophy at the Planck Scale. Cambridge University Press, pp. 213–255.

Bergmann, P.G., Komar, A., 1972. The coordinate group symmetries of general relativity. International Journal of Theoretical Physics 5, 15–28.

Bergmann, P.G., 1961. Observables in general relativity. Reviews of Modern Physics 33 (4), 510–514.

Bergmann, P.G., 1977. Geometry and observables. In: Earman, J., Glymour, C., Stachel, J. (Eds.), Foundations of Space-Time Theories. University of Minnesota Press, pp. 275–280.

Dicke, R.H., 1964. The many faces of Mach. In: Chiu, H.-Y., Hoffmann, W.F. (Eds.), Gravitation and Relativity. W.A. Benjamin, Inc., pp. 121–141.

[25] Hence, we have here an empirical argument for ontic structural realism that evades the standard 'no relations without relata' objection. The relations are the correlations here (the gauge invariant, complete observables), and the 'relata' would be the gauge variant, partial observables. But the partial observables, being gauge variant, do not correspond to physical reality (at least not in any fundamental sense): only the complete observables do. We cannot *decompose* the correlations in an ontological sense, though we clearly can in a epistemic sense—indeed, the correlates constitute our 'access points' to the more fundamental correlations.

Dieks, D. (Ed.), 2006. The Ontology of Spacetime. Elsevier, Amsterdam.
Dittrich, B., 2006. Partial and complete observables for canonical general relativity. Classical and Quantum Gravity 23, 6155–6184.
Earman, J., 1970. Who's afraid of absolute space? Australasian Journal of Philosophy 43, 287–317.
Earman, J., 1989. World Enough and Space-Time. Absolute versus Relational Theories of Space and Time. MIT Press.
Earman, J., 2003. Ode to the constrained Hamiltonian formalism. In: Brading, K., Castellani, E. (Eds.), Symmetries in Physics: Philosophical Reflections. Cambridge University Press, pp. 140–162.
Earman, J., 2006a. The implications of general covariance for the ontology and ideology of spacetime. In: Dieks, D. (Ed.), The Ontology of Spacetime. In: Philosophy and Foundations of Physics, vol. 1. Elsevier, pp. 3–23.
Earman, J., 2006b. Two challenges to the requirement of substantive general covariance. Synthese 148 (2), 443–468.
Eddington, A.S., 1958. The Philosophy of Physical Science. University of Michigan Press.
Einstein, A., 1918a. Dialogue about objections to the theory of relativity. In: Engel, A. (Ed.), The Collected Papers of Albert Einstein, vol. 7, The Berlin Years: Writings: 1918–1921, English Translation of Selected Texts. Princeton University Press, 2002.
Einstein, A., 1918b. On the foundations of the general theory of relativity. In: Engel, A. (Ed.), The Collected Papers of Albert Einstein, vol. 7, The Berlin Years: Writings: 1918–1921, English Translation of Selected Texts. Princeton University Press, 2002, pp. 33–35.
Einstein, A., 1918c. Prinzipielles zur allgemeinen Relativitätstheorie. Annalen der Physik 55, 241–244 (Translated by Julian Barbour).
Einstein, A., 1921. Four lectures on the theory of relativity, held at Princeton university in May 1921. In: Engel, A. (Ed.), The Collected Papers of Albert Einstein, vol. 7, The Berlin Years: Writings: 1918–1921, English Translation of Selected Texts. Princeton University Press, 2002, pp. 261–368.
Einstein, A., 1999. On the ether (1924). In: Saunders, S., Brown, H.R. (Eds.), The Philosophy of Vacuum. Oxford University Press, pp. 13–20.
Giulini, D., 2006. Remarks on the notions of general covariance and background independence. In: Stamatescu, I.O. (Ed.), Approaches to Fundamental Physics: An Assessment of Current Theoretical Ideas. Springer-Verlag.
Kretschmann, E., 1917. Über den physikalischen Sinn der Relativitätspostulate. A. Einsteins neue und seine ursprüngliche Relativitätstheorie. Annalen der Physik 53, 575–614.
Misner, C.W., Thorne, K.S., Wheeler, J.A., 1973. Gravitation. W.H. Freeman and Company.
Montesinos, M., Rovelli, C., Thiemann, T., 1999. An SL(2, $\mathbb{R}$) model of constrained systems with two Hamiltonian constraints. Physical Review D 60, 044009.
Pitts, J.B., 2006. Absolute objects and counterexamples: Jones-Geroch dust, Torretti constant curvature, tetrad-spinor, and scalar density. Studies in the History and Philosophy of Modern Physics 37, 347–371.
Pooley, O., forthcoming. The Reality of Spacetime. Oxford University Press.
Redhead, M.L.G., 1975. Symmetry in intertheory relations. Synthese 32, 77–112.
Rickles, D., 2005a. Interpreting quantum gravity. Studies in the History and Philosophy of Modern Physics 36, 691–715.
Rickles, D., 2005b. A new spin on the hole argument. Studies in the History and Philosophy of Modern Physics 36, 415–434.
Rickles, D., 2006. Time and structure in canonical gravity. In: Rickles, D., French, S., Saatsi, J. (Eds.), The Structural Foundations of Quantum Gravity. Oxford University Press, pp. 152–195.
Rickles, D., 2008. Symmetry, Structure, and Spacetime. Philosophy and Foundations of Physics, vol. 3. Elsevier, Amsterdam.
Rovelli, C., 1997. Halfway through the woods: Contemporary research on space and time. In: Earman, J., Norton, J. (Eds.), The Cosmos of Science. University of Pittsburgh Press, Pittsburgh, pp. 180–223.
Rovelli, C., 2001. Quantum spacetime: What do we know? In: Callender, C., Huggett, N. (Eds.), Physics Meets Philosophy at the Planck Scale. Cambridge University Press, pp. 101–122.
Rovelli, C., 2002. Partial observables. Physical Review D 65, 124013–124013–8.
Rovelli, C., 2003. A dialog on quantum gravity. International Journal of Modern Physics D 12 (9), 1509–1528.

Rovelli, C., 2006. The disappearance of space and time. In: Dieks, D. (Ed.), The Ontology of Spacetime. In: Philosophy and Foundations of Physics, vol. 1. Elsevier, Amsterdam, pp. 25–36.
Rynasiewicz, R., 1996. Absolute versus relational space-time: An outmoded debate. The Journal of Philosophy 45, 407–436.
Sklar, L., 1974. Space, Time, and Spacetime. University of California Press.
Smolin, L., 2003. Time, structure and evolution in cosmology. In: Ashtekar, A., Cohen, R.C., Howard, D., Renn, J., Sarkar, S., Shimony, A. (Eds.), Revisiting the Foundations of Relativistic Physics, Festschrift in honor of John Stachel. In: Boston Studies in the Philosophy of Science, vol. 234. Kluwer Academic, pp. 221–274.
Smolin, L., 2004. Quantum theories of gravity: Results and prospects. In: Barrow, J.D., Davies, P.C.W., Harper, C.L. (Eds.), Science and Ultimate Reality. Cambridge University Press, pp. 492–526.
Smolin, L., 2006. The case for background independence. In: Rickles, D., French, S., Saatsi, J. (Eds.), The Structural Foundations of Quantum Gravity. Oxford University Press, pp. 196–239.
Stachel, J., 1993. The meaning of general covariance: The hole story. In: Earman, J., Janis, A., Massey, G. (Eds.), Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum. University of Pittsburgh Press, Pittsburgh, pp. 129–160.
Torre, C.G., 1993. Gravitational observables and local symmetries. Physical Review D 48, R2373–R2376.
Wheeler, J.A., 1964. Gravitation as geometry—II. In: Chiu, H.-Y., Hoffmann, W.F. (Eds.), Gravitation and Relativity. W.A. Benjamin, Inc., pp. 65–89.

CHAPTER 8

Understanding Indeterminism

Carolyn Brighouse*

Abstract

I examine different varieties of determinism that have been proposed in the context of the hole argument, and describe the basic intuitive idea of determinism. I argue that substantivalists should embrace the indeterminism of the hole argument, because this does most justice to the basic intuitive idea of determinism.

1. INTRODUCTION

In 1987 Earman and Norton published "What Price Substantivalism? The Hole Story" in which they argued that a commitment to substantivalism entails that General Relativity (GR) is indeterministic.[1] The paper led to a flood of articles, and spawned various previously unheard of brands of substantivalism, and various different characterizations of determinism.

Here I want to take yet another look at some of the issues that arise in the hole argument. I am not here to defend any particular brand of substantivalism, but rather to examine some of the varieties of determinism that have been discussed since their paper. I will start with some preliminary remarks about determinism, and describe what I take to be the basic intuitive idea behind the notion of determinism. Next, I will very briefly review the hole argument; I will then examine the responses to the hole argument of Brighouse, Butterfeld, Melia and Skow, and discuss the varieties of determinism that are presupposed by these responses. I will argue that none of these varieties determinism sits well with the basic intuitive

* Department of Philosophy, Occidental College, Los Angeles, USA

[1] The argument applies in principle to all local space-time theories, but can arguably be avoided in space-time theories in which there is "absolute" geometric structure (e.g. Special Relativity where the metric is immutable). Whether it can be so avoided is not something that I will take issue with here; I will rather focus on the argument in the context of General Relativity.

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00008-9

© Elsevier BV
All rights reserved

notion of determinism. And finally I will argue that any substantivalist should embrace the indeterminism of the hole argument, for doing so is the most natural way to accommodate the basic intuitive notion of determinism.

2. A FEW WORDS ABOUT DETERMINISM

Earman and Norton take worlds to be the fundamental bearers of the predicate "deterministic". Others, some of whom I comment on below, have taken laws or theories to be the fundamental bearers of the predicate. I don't think that anything of substance will ultimately hinge on which of these characterizations you prefer. I will follow Earman and Norton here, and when necessary will re-characterise versions of determinism accordingly.

Determinism is a metaphysical fact about a world, and an irreducibly modal one at that. So whether determinism reigns for a world is not just settled by physics alone. It will depend on your ontology, the modal facts at the world, and more generally on the interpretation you give your theory.

It is fairly straightforward to characterize the basic intuition of determinism, and most commentators agree with this basic intuition: A world is deterministic if, given the way the world is now (or was in the past, or is in some region, depending on the kind of determinism we are interested) and the laws of our theory, there is just one way the world can be. Belot characterizes the intuitive idea of determinism thus: "Our world is deterministic if there is only one physically possible future compatible with our past" (Belot, 1995, p. 185).[2]

One might try to argue that this intuition is not the right one underlying our notion of determinism and that we should recognize that sometimes determinism reigns even when there are many physically possible distinct futures given the initial state of the world. Skow does this in Skow (2004), and Brighouse also suggests this as an option in Brighouse (1997). But if you do this then you had better have a convincing case for why sometimes distinct physically possible futures do not violate determinism, while at other times they do. While giving a convincing case for this may be something that can be done, I will argue below that it hasn't yet

[2] This seems to capture the intuitive idea that most have in mind when discussing determinism; thus Butterfield says that a world is deterministic when "A single physically possible world is specified by the facts on a certain region of spacetime" (Butterfield, 1989, p. 12). But we have to be careful even here, for depending on your preferred model for possible worlds the way Butterfield puts this may have different consequences. After all, Lewis thinks no two qualitatively isomorphic possible worlds differ in any way whatsoever, and in particular, no two such worlds differ in what they represent as possibilities for individuals within the worlds. Lewis leaves it up to us to decide if we would prefer to adopt his plurality of worlds in which there are a plurality of qualitatively isomorphic worlds—all the same, call this the liberal plurality, or if we would prefer his more conservative plurality of possible worlds where there is only a single possible world corresponding to each equivalence class of qualitatively isomorphic worlds in the liberal plurality. If we take the conservative plurality then on Butterfield's characterization we are committed to never having qualitatively isomorphic futures violate determinism, and we shall see below that this is bad. If we take the liberal plurality then presumably, on Butterfield's characterization of determinism, determinism can never be secured no matter what your theory or world. For you'll never have a unique possible world picked out by the state of the world at a time (or in the past, or on some region); at best you will be able to determine an equivalence class of qualitatively isomorphic worlds. Of course, we can re-characterize determinism for the liberal plurality so as not to have this consequence, but I think it would better at the outset to try to spell out the intuitive idea of determinism in such a way that presupposes as little baggage as possible, and then worry about how to fit it within ones preferred metaphysical framework. By talking about different physically possible futures compatible with our past, I think Belot's characterization above hits this mark.

been done in a fully defensible manner. Ultimately I will suggest that the appropriate approach to characterizing determinism is simply to accept the basic intuitive idea.

3. THE INDETERMINISM OF THE HOLE ARGUMENT

The charge of indeterminism in the hole argument arises from the fact that GR is a generally covariant theory. Given a model $\langle M, g \rangle$, any model diffeomorphic to it is a model of the theory. The diffeomorphism that relates a pair of models can be the identity map except in some region "the hole", in which case we have two models that are identical throughout space-time except in the hole, where they diverge.

If we take these models to represent distinct physical possibilities, and we endorse the basic idea of determinism described above, where determinism reigns if there is only one physically possible future compatible with our past or present, then determinism is violated by a pair of hole diffeomorphs. According to Earman and Norton a substantivalist—or more precisely a manifold substantivalist—who takes the points of the manifold to represent points of space-time independently of the geometric object fields defined at them, will view these models as representing distinct physical possibilities, and so we get the charge that the manifold substantivalist will take the worlds described by these models of GR to be indeterministic.

The differences between a pair of hole diffeomorphism related models, or between the physically distinct possibilities that for the manifold substantivalist they purportedly represent, are subtle ones. The possible worlds described by these models are globally qualitatively isomorphic, but they differ in the way geometric properties and matter fields are distributed over the points of space-time in the hole. Within the holes two such worlds differ over which spacetime points are playing which role.

This feature has led some (Butterfield, Brighouse, Melia and Skow) to think that the violation of determinism here is only apparent, and that the correct analysis of the hole argument will be one that shows why the hole diffeomorphs don't count as determinism violators. People have done this in different ways. The earliest moves were motivated by counterpart theory and endorsed a version of determinism on which no pair of qualitatively isomorphic worlds when taken by themselves would constitute a violation of determinism (Butterfield and Brighouse).[3] More recently Melia and Skow have endorsed more sophisticated positions in which it is argued that while hole diffeomorphism related worlds do not, by themselves, count as determinism violators, there nevertheless are some pairs of qualitatively isomorphic worlds that, when taken by themselves, do.[4]

[3] These early positions carried with them, a rejection of manifold substantivalism arguing either that at most one of a diffeomorphism related class of models represented a possible world (Butterfield), or that all such models represented the same possible world (Brighouse). They then held true to the intuitive idea of determinism.

[4] One might wonder what is meant by talking about a pair of worlds failing to constitute a violation of determinism when taken by themselves. What I mean is, given a particular pair W and W', whether by just considering W and W' (and no other worlds described by the theory), we can conclude that we have a violation of determinism. Of course, whether any given world, W, actually is indeterministic will depend on whether there are any other worlds represented by models of the theory that are duplicates of W in the past but not global duplicates of W.

4. DEFINING DETERMINISM: EARLY ATTEMPTS AT AVOIDING THE HOLE ARGUMENT

Butterfield and Brighouse urged that one adopt a David Lewis style analysis of determinism. According to David Lewis we should take counterpart relations to depend on qualitative similarity, and so our definition of determinism should likewise be sensitive only to differences between worlds that are qualitative. The best articulation of this kind of program for defining determinism can be found in Belot (1995). Borrowing from Lewis (1986) we can talk of duplicates: "two things are duplicates iff (1) they have exactly the same perfectly natural properties and (2) their parts can be put into correspondence in such a way that corresponding parts have exactly the same perfectly natural properties and stand in the same perfectly natural relations" (Lewis, 1986, p. 186). Perfectly natural properties are those that "carve nature at the joints and *only* at the joints" (Belot, 1995, p. 188), and we get a sense of what these are from our best physical theories. Following Lewis, Butterfield and Brighouse argued that we should adopt the following definition of determinism:

DEFINITION 1. W is deterministic if, whenever W' is physically possible relative to W and t, t' and $f : W_t \rightarrow W'_{t'}$ are such that f is a duplication, there is some duplication $g : W \rightarrow W'$ (Belot, 1995, p. 190).

Intuitively the idea here is that if a world W is such that all those worlds that have duplicate initial segments to W are global duplicates of W then W is deterministic. And this is true, according to Definition 1, even if the initial duplication, f, between W and another world, W', happens not to be extendible to a global duplication. For one might have worlds that exhibit enough symmetries up to some time, but not after that time, such that some duplications of initial segments are not extendible to a global duplication, but, the defender of Definition 1 would claim, if the worlds themselves are global duplicates this should not matter to determinism. According to Definition 1 then, if all worlds that are initial duplicates of W turn out to be global duplicates of W, W's determinism is secured.[5]

Obviously the models of the hole argument no longer count as determinism violators, since whenever we have duplicate initial segments of the physically possible worlds represented by hole diffeomorphs the worlds are global duplicates. But Definition 1 has been criticized precisely because it fails to take any pairs of worlds that are global duplicates as constituting, by themselves, a violation of determinism. It has been argued by Belot, and following him, Melia and Skow, that there are cases of worlds that are deterministic according to Definition 1 but for which we have very strong intuitions of indeterminism.

The most compelling such example is the tower collapse example first discussed by Mark Wilson (1993): Imagine a world empty but for Newtonian spacetime and a spherically symmetric homogeneous planet with nothing but a cylin-

[5] Definition 1 deals with "future" determinism—dependence of the future on the past. One could clearly characterize past determinism or other forms of determinism in the spirit of this definition. The central claim is that if all local duplicates, for a specified region of interest R, are global duplicates, then determinism is secured.

drically symmetric homogeneous tower with a cone traveling through space in such a way that it's apex will hit the top of the tower in the center. The laws decree that the tower will buckle in one direction, but while the shape of the buckle is determined, the direction of the buckle is not. We can imagine spherically symmetric worlds in which this takes place. But we want to be able to say that two such worlds W and W' differ in that the tower could collapse in a different direction in the two worlds despite the fact that the worlds are global duplicates.

If we find this example compelling, then we are forced to give up Definition 1. Should we find this example to be a compelling case of indeterminism? Yes. There seem to be genuinely different physical possibilities for the tower. There are many different parts or regions of the planet that the tower could have landed on, but in any given world it landed on just one, some of the other physically possible worlds are ones in which it lands on a different region. To be sure for any initial duplication between the worlds there is a global one that matches the whole worlds qualitatively, but this does not seem to be enough for a judgment of determinism. For, despite this global duplication between worlds, which region of the planet the tower collapses on is simply undetermined by the laws and initial conditions. If it were determined then it would have to be the same region in all worlds that are initial duplicates, but there is nothing in the laws or the state of the world that would serve to determine which region this is. And it would be absurd to maintain that it is determined but we don't and can't know which region it is: there simply are no facts about the laws and the state of the world before collapse that **could** determine the direction in which the tower will fall.

Melia gives a convincing assessment of the indeterminism in the tower collapse case: what the Lewis definition masks by requiring only that there exists a global duplication between the worlds given that there is an initial duplication is the fact that parts of the worlds can evolve in different ways even when the worlds as a whole are duplicates. In fact he suggests that we have a recipe for generating this kind of indeterminism: "We imagine a world that contains a collection of objects which share not only their intrinsic qualitative properties at time t but also their relational ones as well. Suppose that, at some time t^* later than t, one of these duplicate objects changes one of their intrinsic or extrinsic properties." (Melia, 1999, p. 650), in any such case we will have a violation of determinism just like the one in the tower example. And he provides some examples of the application of this recipe: Imagine a world empty but for a circle of bald philosophers, and a law that says that one will grow a hair: surely, given the symmetry of the situation any one of them could be the lucky one (Melia, 1999, p. 650). And so it seems that Definition 1 must be rejected.

5. DEFINING DETERMINISM: DRIVING THE WEDGE BETWEEN THE HOLE AND THE TOWER

So the tower collapse example seems to provide a compelling case of indeterminism and accordingly we need to replace Definition 1. If we want to maintain that the worlds of hole argument do not represent genuine cases of indeterminism,

while the tower world is genuinely indeterministic, then we need a definition that can judge these cases differently. Of course, not just any account that accommodates these intuitions would thereby be acceptable even to those who have these intuitions. Such an account must come with some justification of its adequacy. So when assessing whether we have the correct definition to replace Definition 1 we must not only assess whether it saves our "tower" intuitions on the one hand and our "hole" intuitions on the other, but we had better assess whether it saves these intuitions in a way that can be well motivated. In this section and the next I will describe the best attempt to save these intuitions that I know of, and then in Sections 7 and 8 I will argue that this attempt is not defensible.

Belot, Melia and Skow all share the intuition that the tower example is indeterministic, and all agree that Definition 1 needs to be rejected. Each offer replacements that are essentially equivalent to one another (Belot only to go on to reject this kind of replacement as inadequate), but in addition, as part of the motivation for their replacements for Definition 1, Melia and Skow present a detailed account of what went astray in the hole argument.

According to Melia the apparent indeterminism in the hole argument arises because "The space-time point **a** which actually plays role F at t^* and the space-time point **b** which actually plays the role G could have played each other's roles" and immediately here he asks us to "Note that **a** and **b** do not exist yet at t—they come to exist at t^*" and he goes on to say "thus the features of space-time that play an essential role in Earman and Norton's demonstration of [in]determinism are these: (a) the fact that new space-time points come into existence [and] (b) the fact that it is possible for two future space-time points to have had each other's properties" (p. 646).

Melia argues that if we endorse these two assumptions we will rule indeterministic many theories that are intuitively deterministic. His example of such a theory is a relationist theory (that describes worlds with no space-time points) in which there are two types of particles P^+ particles and P^- particles. After five minutes P^+ particles decay into P^- particles. The relative configuration of the P^- particles is same as the relative configuration of the P^+ particles at the earlier time. Compatible with this theory there are apparently different physically possible, but globally duplicate worlds. We can imagine a world in which there are two P^+ particles **a** and **b**. Later there will be two P^- particles **c** and **d**. One scenario is that **a** decays into **c** and **b** decays into **d**, but it is perfectly compatible with the theory that **a** decay into **d** and **b** decay into **c**. So we have another example of indeterminism. It is not determined which P^- particle is the product of the decay of a given P^+ particle. These worlds are thus indeterministic. But, according to Melia, this is absurd. Melia thinks that the P^+/P^- worlds are as deterministic as any worlds one could ever wish to see. According to Melia we should not be so quick to judge worlds as indeterministic when the differences between their futures consist solely of objects that didn't exist in the past playing each other's roles in the futures.

So the culprits in P^+/P^- worlds and in the hole argument according to Melia are our assumptions (a) and (b). His intuition is that you don't want determinism to fail in worlds that are qualitatively isomorphic to one another just because there are entities that don't exist in the duplicate initial parts of the world, which appar-

ently play each other's roles in the future. It is precisely this that is going on in the hole argument and in the offending theory of P^+ particles. And clearly this is not what's going on in the Tower collapse case. There you have no things that pop into existence in the future to cause a violation of determinism, instead all the objects that exist (the tower, the planet, the points of space, the cone), exist throughout time.

So here we have a principled difference between cases of the "hole" type on the one hand and cases of the "tower" type on the other, and this gives us one way to settle matters of determinism. If you are faced with a purported violation of determinism, check to see of the indeterminism arises because entities that didn't exist in the past are playing each other's roles in the purportedly different futures. If they are, and that is the only difference in the futures, then it is not a genuine case of indeterminism, if they are not, or there are other differences in the futures then it is a case of genuine indeterminism. Perhaps this principled distinction will be one on which to base a better definition of determinism than Definition 1.

In fact this criterion for distinguishing between hole argument-type examples and tower collapse-type examples also allows us to classify some other examples in the literature: Gordon Belot has two examples (Belot, 1995), that he uses for purposes different from those of Melia, which allow one to further see how Melia's approach works; I'll call them PAIR and CONTINUUM.

PAIR: Imagine the world PAIR which is a relationist world empty but for the existence of a single alpha particle, A. The laws decree that after 13 years A will decay into two beta particles, B_1 and B_2. According to the laws the time of decay is determined, the products of decay and all their physical magnitudes are determined, and the angle between the products of decay is determined.

CONTINUUM: Imagine the world CONTINUUM that has Newtonian space-time structure (with points of space that exist through time). In such a world exists a single alpha particle, which decays after 13 years, but in this world alpha particles decay into continuum many B particles. The B particles form a spherical shell around the place of decay and each particle moves at same rate along its radius.

Belot originally had the intuition that both of these worlds were indeterministic. Paraphrasing Belot, in PAIR before the decay there seem to be two possibilities for the future of the world corresponding to two different extensions of the initial duplication of worlds W and W'. Let t, t' and f be such that $f : W \to W'$ is a duplication. If we restrict attention to $t < 13$ years, and label the particles in W and W' as A, B_1, B_2 and $A'B'_1, B2'$ then there are two different ways to extend f. We could extend the initial duplication by mapping $B1$ onto $B1'$ and $B2$ onto $B2'$, or alternatively by mapping $B1$ onto $B2'$ and $B2$ onto $B1'$. This second extension seems to license the claim that $B1$ could have been $B2$, and Belot continues by saying "If we were somehow observing W, we could say that although as it turns out, $B1$ is *this* particle, and $B2$ *that* one, it could have been the other way round. I would say that W is indeterministic: before the decay, there are two possibilities for its future." (Belot, 1995, p. 192).

Once we think that PAIR is indeterministic then this suggests that CONTINUUM is indeterministic also. Again, paraphrasing Belot, if we focus in CONTINUUM on two particles $B1$ and $B2$, while there is only one extension of the initial duplication to a global duplication, still we can consider a rotation that maps the world line of $B1$ onto the world line of $B2$, and while this is not an extension of the duplication since it doesn't preserve distance between the particles and points of space, it does seem to license the claim that after all B_1 could have been B_2, or at least B_1 could have had B_2's trajectory and vice versa.[6]

But despite this these seem like particularly strange and subtle violations of determinism. Just try to state the facts that are undetermined in PAIR's future and you'll see it is hard to do. In PAIR what is undetermined in the future is whether a currently non-existing particle, $B1$, is $B1$ or $B2$. It is certainly not obvious that this makes sense and you can understand why someone might be motivated to try to reject this reasoning and argue that these examples are not really cases of indeterminism. Melia's approach allows us to do just that.

In both of the examples of PAIR and CONTINUUM we are dealing with worlds that are global duplicates, and they have duplicate initial segments up to time t ($= 13$ years), but they are worlds in which particles that apparently violate determinism only pop into existence after the initial duplicate segments. In fact they exhibit precisely the kind of indeterminism that Melia was so appalled by in the $P+$ world: in both PAIR and CONTINUUM the apparently different futures differ in virtue of objects that didn't exist in the past playing each other's roles in the different futures. And so, provided we have the intuition that these worlds are deterministic, which presumably Melia has, then perhaps this principled distinction between the hole case and the tower case has something going for it, and is something that will motivate a replacement for Definition 1.

Skow (2004) provides further support for this kind of assessment of what is to blame for the alleged indeterminism in the hole argument. He considers the Laplacean conception of determinism where, in a deterministic world, Laplace's demon, who knows all the information about the present as well as the laws of nature, would be able to deduce all the information about other times. And he (Skow) asks what we should reasonably expect the Laplacean demon to be able to predict: is it all qualitative information from qualitative information? Is it all qualitative and non-qualitative information from qualitative and non-qualitative information? Or, is it something in between? Skow urges that it is unreasonable to expect the demon to be able to predict non-qualitative information about future objects that don't exist at the time about which he has all qualitative and non-qualitative information. If you insist that he predict non-qualitative information about the future at all, it should simply be information about objects that existed, and had names prior to the relevant time. To expect more would be tantamount to expecting him to give proofs of substantive sentences in first order logic that contain names that don't occur in the premises—it can't be done!

[6] And in fact this intuition led Belot to think that it was not clear that a counterpart theorist could arrive at an account of determinism that would rescue the substantivalist from the threat of indeterminism in the hole argument.

6. SOPHISTICATED DETERMINISM

Melia and Skow, motivated by the kinds of considerations above, offer new definitions of determinism to replace Definition 1. Melia offers two slightly different accounts while Skow offers one that is essentially equivalent to Melia's second.

Melia's first resolution: Take any two worlds, W and W' described by your theory T, take any two histories up until some time t, h in W and h' in W', and take any two objects o from h and o' from h'. The worlds W and W' are deterministic if and only if, if h and h' are duplicates and o and o' are duplicates (in the sense of instantiating precisely the same intrinsic and relational qualitative properties) then W and W' are duplicates and o and o' are duplicates for all time.[7]

This captures the idea that indeterminism arises when objects that shared all their intrinsic and relational properties in the past fail to share them in the future even when the worlds taken as a whole are duplicates. It also respects the distinction between the worlds of the hole argument, PAIR, CONTINUUM, and $P+$ decay on the one hand and the worlds of the tower collapse on the other, for in all of these former cases the futures contain objects that only came into existence after the initial segments. In each of these cases all the objects that were in the initial segments that initially had the same qualitative properties go on to have the same qualitative properties. So according to the first resolution the determinism of all these worlds is secure.

Melia's second resolution/Skow's resolution: The intuitive idea behind this resolution is that if we are dealing with two worlds that are initial duplicates then for the world to be deterministic that duplication must be extendible to a global duplication of the worlds.[8]

W is deterministic if, whenever W' is physically possible with respect to W and t and t' and $f : W_t \to W'_{t'}$ are such that f is a duplication there is some duplication $g : W \to W'$ whose restriction to W_t is f.

Just as the first resolution drives the wedge between the hole and the tower, so too does this resolution. PAIR and CONTINUUM come out as deterministic,[9] as do the worlds of the hole argument, for the hole diffeomorphs themselves provide the extension of the initial duplication. $P+$ decay worlds also get ruled deterministic. And according to the second resolution the tower case is indeterministic: there are some initial duplications that are not extendible to global duplications, and these are precisely those that seem to license the claim that the tower could have collapsed in a different direction.

[7] Melia formulates determinism here in terms of theories rather than worlds, for uniformity I am adapting it to an account in terms of worlds. Melia's own account can be found on p. 654 of Melia (1999).

[8] Melia and Skow give different wordings of the definition, but both are equivalent to Definition 2 of Belot's which can be found in Belot (1995, p. 191). Rather than giving these different equivalent versions here, since both Skow and Melia claim that they are endorsing Belot's Definition 2, I will give Belot's version.

[9] In PAIR we have two extensions of the initial duplication whose restrictions to W_{13} are f, and in CONTINUUM we have a unique extension whose restriction to W_{13} is f. See Belot (1995, Section 5).

In fact we can see that the first and second resolution will rule out precisely two kinds of candidate violations of determinism in globally duplicate worlds as counting as genuine violations of determinism. These two kinds of candidate violations correspond precisely with features (a) and (b) above that according to Melia caused us to think that the apparent violation of determinism in the hole argument was genuine. Remember, feature (a) was the fact that new space-time points come into existence and feature (b) was the fact that it is possible for two future space-time points to have had each other's properties. So lets call the corresponding candidate violations of determinism violations of kind *A* and kind *B*.

Candidate violations of determinism of kind *A* occur when objects are created. According to these resolutions no collection of globally duplicate worlds will ever be deemed as violating determinism simply because of object creation. For in such cases according to the first resolution the only objects that are relevant to the question of determinism are the ones that exist in or at the region in the antecedent of the condition for determinism, i.e., if we are thinking in terms of the past determining the future, the only objects that have to be duplicates in the futures of these worlds are the ones that actually existed in the past and were duplicates there. So if you have two objects created in the futures that are qualitatively different, the fact, if it is a fact, that all previously existing duplicate objects remain duplicates and the fact that the worlds taken as a whole are global duplicates are sufficient to secure determinism.[10] And the second resolution has this consequence too: to secure determinism it has to be that the initial duplication between worlds is extendible to a global duplication, it's what that duplication does to objects that existed in the past that is heavily constrained. Beyond the fact that the duplication has to be global there are no constraints on what the duplication does to objects that exist only in the future, so even if there are many different extensions of the initial duplication to a global duplication the world is still deemed deterministic.

The second kind of candidate violation of determinism, kind *B*, involves objects playing each other's roles. While globally duplicate worlds can violate determinism simply in virtue of objects playing each other's roles, as evidenced by the tower case in which a particular part of the planet could have been playing the role of a different part, or a particular point of space at the elbow of collapse could have been playing the role of another point of space, nevertheless, according to the two resolutions, playing each other's roles in duplicate worlds only leads to a violation of determinism when those roles are qualitatively different. Thus, in resolution one, any two objects that are initial duplicates must remain duplicates, and this is sufficient in duplicate worlds to secure determinism, even if we think it

[10] Imagine an initially empty relationist world (let's say in which we can talk about creation of the world by God at $t = 0$) in which a law decrees that after 13 years two qualitatively identical black particles are created 10 feet apart. Such a world is deterministic. So too is a similar empty relationist world in which a law decrees that after 13 years two particles are created and one is red and one is blue. As long as the worlds are qualitative duplicates it doesn't matter whether all created particles are duplicates or not for determinism to be secured—created particles don't matter. Being empty initially is not crucial to these examples, but the worlds in the examples here do need to exhibit a high degree of symmetry. For the Melia/Skow resolutions to get some indeterminism qualifying traction there need to be objects that exist prior to the particle creation that could have different qualitative properties or relations after the particle creation. Obviously this can't happen in empty worlds, but neither can it happen in worlds that exhibit enough symmetries. So imagine a relationist world in which there is a law of red/blue particle creation like that above, but in which there is a single object, *P*, existing throughout time and the particles are created equidistant from and on opposite sides of *P*, so that after creation the red particle, *P* and the blue particle are collinear. This world is deterministic according to Melia and Skow.

is true that these objects could have played each other's roles in the future. And in resolution two, the fact that the initial duplication is extendible to a global duplication means that duplicates in the past must remain duplicates. So even if we could give an argument that two futures could be different in virtue of which object is playing which role if these roles are duplicate roles these resolutions will never see those distinct futures as violating determinism.

Resolutions 1 and 2 do come at something of a price over the notion of determinism championed by Lewis and endorsed by Butterfield and Brighouse, but I it is a price that has to be paid by anyone who wants to preserve the intuitions of the tower collapse example, and Melia and Skow are perfectly upfront about their willingness to pay this price. We are now countenancing haecceitistic differences between worlds and allowing that these haecceitistic differences at least sometimes count as a violation of determinism. For the worlds of the tower example are globally qualitative isomorphic but different nonetheless in terms of which objects are playing which. In the discussion below I will not question this commitment to haecceitistic differences, for I think the intuition that the tower example is indeterministic is a very strong one. Those who find haecceitistic indeterminism too high a price to pay will have to reexamine their intuitions about the tower example.

At this point the reader might be thinking—well so much the better for resolutions one and two. They provide a nice way to avoid the charge of indeterminism in the hole argument, and they allow us to maintain that the tower collapse world is indeterministic, and they provide a way to dismiss the claims of indeterminism in the artificial examples of $P+$, PAIR and CONTINUUM. But I am not so sure that these judgments about candidate violations of determinism of kinds A and kinds B are ultimately defensible, and it is to this I turn next.

7. IS SOPHISTICATED DETERMINISM RIGHT?

We have two characterizations of determinism that judge examples in a way that accords with intuitions that at least some of us have: the hole argument just doesn't seem to show that GR is indeterministic, while the tower case sure does look like a case of indeterminism. But fitting with these intuitions does not by itself provide an argument for adopting resolutions one or two. Moreover, it seems to me that these resolutions are not ultimately defensible: They tie the issue of determinism too closely to A and B above. I think that A is simply a red herring; that there is absolutely no reason to think that apparent violations of determinism that feature object creation should not count as genuine violations of determinism. And I think one can motivate an intuition that contrary to B, if we allow futures that differ simply in virtue of which objects play which roles to count as violating determinism then we should (at least sometimes) be willing to countenance such violations even when the roles those objects are playing are duplicate roles. Here I will start to try to flesh out some intuitions that motivate the idea that sometimes candidate violations of determinism of kinds A and B are genuine violations, and in the next section I will try to spell out some arguments for these conclusions.

Why be suspicious of sophisticated determinism? Well, surely sometimes particles coming into existence at a later time can give rise to indeterministic developments of worlds. After all, space invaders from infinity in Newtonian space-time seem to threaten determinism there, and naked singularities spewing forth garbage into space-time would threaten determinism in GR. In both of these cases it *seems* to be precisely objects coming into existence in the future that makes it impossible for the laws plus the state of the world in the past to determine the future. But Melia and Skow can accommodate these cases for, for them, in worlds that are initial duplicates but not global duplicates you can have as much particle creation as you like and the worlds will still be indeterministic. This is not because of particle creation exactly, but because the worlds are simply not global duplicates. Take the case of space invaders from infinity: two such worlds may not remain duplicates because some particle came into existence, but according to Melia and Skow it is the properties of previously existing objects that are relevant for determinism, not the fact that a particle came into existence. In this case it is because two previously existing points of space that were duplicates in the past do not remain duplicates when one becomes occupied by a space invader and the other doesn't.[11]

Here's an attempt at characterising an intuition that neither *A* nor *B* seem to be right. Consider CONTINUUM (remember here we have points of space that exist through time with full Newtonian structure and we have a single alpha particle that decays creating continuum many *B* particles), and now consider God thinking of the ways He might populate the world. Before the decay doesn't He have many (continuum many) options as to where to place the *B* particles? He could place B_1 at *this* point of space or *that* one immediately after the decay, and these would be different creative acts, just as in the tower case He could have made the tower collapse in one way rather than another. In both the tower case and the continuum case He had a bunch of options, and failing any law of the theory that forces His hand, the fact that He had a bunch of options seems to suggest that we have indeterminism, at least on the intuitive notion of determinism that we started with here where determinism is violated if there is more than one future compatible with the past. And there isn't anything in the theory that could force His hand—He can throw the dice and place the elbow of collapse in the tower world, or B_1 in CONTINUUM, at any point of space at the relevant time that he pleases. We agree that this is a failure of determinism in the tower case. Surely this is a failure of determinism even in CONTINUUM: don't we have more than one distinct physical possibility for the future given the state of the world at an earlier time?

Consider modifying continuum such that one of the beta particles, say B_{700} is red and all the others are blue. Here, for sure, we have a violation of determinism, since now the example is exactly analogous to the tower. And of course Melia and

[11] The case of naked singularities is a bit different because points of space-time in GR exist only at one point of time—this is what allows Melia and Skow to use their resolutions to escape the indeterminism of the hole argument. So here it is not a case of two previously existing duplicate objects having different properties after some time that would make nakedly singular spacetimes indeterministic, it must be just because the worlds corresponding to these spacetimes simply fail to be global duplicates (the singularities spew forth qualitatively different things in different worlds). In fact, for Melia and Skow, if we had a theory all of whose models describe worlds that are global duplicates, in which the fundamental space-time background consists of points of space-time rather than points of space through time, such a theory could not fail to be deterministic.

Skow can accommodate this as a violation, for in the future there are regions of space-time that could have different qualitative properties: two duplicate parts of the past (regions of space prior to t) sometimes fail to remain duplicates in the future. But why? Well, precisely because those individual B particles could traverse the different trajectories. To adopt the metaphor of the previous paragraph: precisely because there is nothing in the laws constraining God to place any one of those Beta particles on one path over another. But this seems to be independent of whether all the Beta particles are qualitatively identical or not. After all in this case God can decide on a whim to place B_{700} somewhere, but He still has the task of placing all the other particles somewhere too, and this task seems to be essentially the same as his task in CONTINUUM. But doesn't this seem to suggest that if you count this world as indeterministic you should also count CONTINUUM as indeterministic?

And further—consider the state of the world of CONTINUUM immediately after the decay. Of course we recognize that this particle could have traversed that particle's trajectory. In fact it is more concrete than that, after all, all these particles have a different velocity albeit with the same magnitude, and each one of them could have had the velocity of one of the others—the laws of the theory sure don't dictate which they had. And B_1 in Continuum could have had B_{700}'s velocity, just as much as B_1 in the modified example where B_{700} is red could have had B_{700}'s velocity—what the possible velocities for B_1 are seems to be independent of B_{700}'s (or, for that matter, any B particle's) colour.

But these are just fleshings out of intuitions that I have. They are not an argument exactly, but they lead me to think that the resolutions endorsed by Melia and seconded by Skow might drive the wedge between the tower and the hole for the wrong reason.

8. ARGUING AGAINST SOPHISTICATED DETERMINISM

The issues at stake in violations of determinism of kinds A (object creation) and B (objects playing each other's roles) are independent of one another, and so I will address them separately. In each case I am going to argue that sophisticated determinism ends up violating the spirit of the intuitive conception of determinism that we started with in Section 2, namely, that a world is deterministic if and only if there is at most one possible future compatible with its past. Worlds for which there is more than one possible future compatible with the past are indeterministic. For each of the two kinds of candidate violations of determinism A and B I will provide two examples. One of these examples will be classified as deterministic according to sophisticated determinism while the other will be classified at indeterministic. I will argue that the examples should be judged similarly with respect to determinism, and then explain why both examples violate the intuitive conception of determinism and so should be judged as indeterministic.[12] First let's try to factor out B type issues, namely the issue of whether futures may differ in virtue

[12] I should issue a warning here: the examples I am going to consider may seem physically unrealistic; more on this later.

of objects playing each other's roles when those roles are duplicates, and focus on the issue of particle creation.

8.1 Indeterminism of kind *A*: Particle creation

Consider two kinds of collections of worlds that are each somewhat like PAIR.

PAIR-red-blue: This is a relationist world in which initially a single alpha particle, A, exists. The laws decree that after 13 years A will decay into two beta particles. According to the laws the time of decay is determined, the products of decay are determined, and the angle between the products of decay is determined. But one of the beta particles is red and the other is blue, and while the laws decree that one will be red and the other blue the laws are silent on which is red and which is blue.

Now consider the world **PAIR-red-blue-no-decay**: This is a relationist world, but in this world alpha particles never decay, and they have a rather special property of being able to co-exist with each other in the same region of (relationist) space-time for periods of time. Imagine such a world in which initially two co-located alpha particles exist. The laws decree that two (same coloured) alpha particles can co-locate for exactly 13 years, at which point they will move apart from one another, one will turn blue and the other will turn red, and while the angle of separation of the alpha particles after 13 years is determined, which one turns red and which one turns blue is undetermined.

PAIR-red-blue-no-decay is an odd world. We don't typically think that things can co-locate, but I don't think this notion of co-location is incoherent: fields being defined in the same region of space-time do not disturb us. So try, if you can, to bear with me for this example. I will come back the issue of its oddness later.

I think that it is fairly easy to motivate the idea that PAIR-red-blue-no-decay is indeterministic once we have, as I am assuming we have, granted that some classes of duplicate worlds are indeterministic.[13] Afterall, given the state of the world prior to 13 years there are two different ways the future could be, it could be that A_1 is blue and A_2 is red, or it could be that A_2 is blue and A_1 is red. The future is undetermined given the history of the world. And, according to Melia's and Skow's sophisticated determinism, this is precisely the kind of indeterminism (that accords exactly with Melia's indeterminism generating recipe) that arises in duplicate worlds: you have parts of the world that are initial duplicates that do not remain duplicates. And so in fact according to both of their resolutions PAIR-red-blue-no-decay is indeterministic. According to sophisticated determinism however, while PAIR-red-blue-no-decay is indeterministic, PAIR-red-blue is deterministic. For in PAIR-red-blue there are no parts of the world that are initial duplicates that fail to be duplicates in the future. The beta particles that are not duplicates in the futures didn't exist in the past. But why should this matter? It seems to me that PAIR-red-blue worlds have exactly the same kind of indeterminacy after 13 years as PAIR-red-blue-no-decay worlds. In each of these worlds there

[13] If you grant that the tower collapse world is indeterministic you have granted that some classes of duplicate worlds violate determinism.

are futures that contain a pair of particles whose roles could have been switched, and if those roles had been switched the future would have been a different one. And in neither of these cases is there anything about the state of the world in the past and the laws that fixes which particle plays which role in the future. In both PAIR-red-blue-no-decay and PAIR-red-blue we have more than one possible future compatible with a single past, and thus both worlds violate the intuitive conception of determinism. If this indeterminacy is sufficient for a judgment of indeterminism in PAIR-red-blue-no-decay, then why is it not likewise sufficient for a judgment of indeterminism in PAIR-red-blue? Why does it matter whether those particles that could have played each other's roles existed in the past or were the products of decay? For whether they are the products of decay or not in each case (PAIR-red-blue-no-decay and PAIR-red-blue) the worlds are apparently such that there are two possible futures, differing just in terms of which objects play which roles, and the present state of the world doesn't determine which of the two futures will be realized, and it is precisely this kind of situation that violates the intuitive conception of determinism that we started with here. So I would submit that with respect to the issue of determinism there is no relevant difference between PAIR-red-blue and PAIR-red-blue-no-decay. Particle creation here is a red herring. We should accept that violations of determinism can occur where futures differ solely in virtue of created particles having different properties in the different worlds even when those worlds are global duplicates.[14]

Skow does provide some explicit motivation for rejecting the intuitive account of determinism and accepting something like resolution one or two. He points out the unreasonableness of the expectation that the Laplacean demon have to determine nonqualitative facts about the future given qualitative and nonqualitative facts about the past, when those nonqualitative facts about the future involve facts about objects that didn't exist in the past. And, yes, it does seem unreasonable to expect this.[15] But let's think about the demon's predicament in PAIR-red-blue and PAIR red-blue-no-decay: In neither can he determine where the relevant particles will be after 13 years. But why should we think his failure in PAIR-red-blue is more troublesome than his failure in PAIR-red-blue-no-decay. He knows that in PAIR-red-blue there will be two particles in the future, and that there are different options for the color for any given one of these particles. Does it really matter that he doesn't have names for them? Why don't we let him, when he is in this kind of situation, pick names for the future existents—and if, given that he picked names, there is only one way that the future could be, then let this count as him determining the future. Of course most of the time such worlds will come out as

[14] Another example that motivates the same conclusion, that is perhaps no less odd (though odd in a different way; it doesn't require the notion of co-location) is to consider modifying the bald philosopher's case. Consider relationist world PHILOSOPHER in which a circle of bald philosophers sits, and one eventually grows a hair. For Melia this world is indeterministic. But now consider a different relationist world PHILOSOPHER-DECAY, in which a circle of bald philosophers sits and after 13 years they decay into physicists, one with hair and the rest bald. For Melia this world is deterministic. It seems to me that these two examples are not relevantly dissimilar when it comes to the question of determinism. Decay just seems to be a red herring.

[15] But remember, when we confront Laplace's demon with a determinism violating world the prediction task we expect of him is always an unreasonable one, so we shouldn't be too quick to infer from the unreasonableness of a Laplacean prediction task to the conclusion that we are confronting merely an apparent rather than a genuine violation of determinism.

indeterministic—but not always: certainly worlds could come out as deterministic when only one thing comes into existence in the future.

8.2 Indeterminism of kind *B*: Objects playing qualitatively duplicate roles

Let's now turn to the issue of whether we ever have genuine violations of determinism when futures differ just in terms of which objects are playing which roles, where those roles are qualitative duplicates of each other. Remember, Melia and Skow are willing to countenance examples as genuine violations of determinism when futures differ just in terms of which objects are playing which roles, the tower world is just such a case, but for them the only way such violations come out as genuine is when the roles that the objects could have been playing are qualitatively different from the roles they are actually playing. I think one can motivate the view that determinism is violated even when futures differ just in terms of which objects are playing which qualitatively identical roles.

Let's consider a new example:

CONTINUUM-no-decay. This is a lot like the world CONTINUUM: we have points of space through time with full Newtonian structure. And we have a world empty but for continuum many co-located alpha particles. Alpha particles are special in that they can be co-located for exactly 13 years but no longer, at which point the laws decree that they will start to move away from the point of co-location, and will form a spherical shell around the point of co-location each moving at the same rate along its radius. But while the laws decree that each of the possible trajectories after 13 years will be occupied by an alpha particle the trajectory of any one of the alpha particles is undetermined.

According to Melia and Skow this world is deterministic. But now compare this with the following world:

Continuum-no-decay-red, which is just like CONTINUUM-no-decay but in which the laws decree that one of the alpha particles after 13 years will be red. Clearly CONTINUUM-no-decay-red is indeterministic on Melia's and Skow's accounts, for (according to resolution one) we have parts of the world that are duplicates initially, but they are not duplicates in the future (pick the alpha particle that becomes red and any other alpha particle, these are duplicates for the first thirteen years, but not after thirteen years), and (according to resolution two) any initial duplication that maps the afore mentioned two objects onto one another fails to be extendible to a global duplication. Moreover, this example is very similar to the example of the Tower world. Here, just as there, we have points of space through time with full Newtonian structure. In both cases we have worlds that are completely cylindrically symmetric for an initial period of time, and then after that time the symmetry is broken. So it would seem that anyone who thinks that the tower world is indeterministic should also think that this world is indeterministic.

Let's think carefully about why we should judge this world to be indeterministic. In CONTINUUM-no-decay-red, just as in the tower world, it would seem that we have multiple different possible futures compatible with a single history, and this is so because the same kinds of counterfactual claims seem to be true in each case. In the tower world we think that there are many possible buckle directions for the tower. And we think this is true because we think that some relevant counterfactuals are true. What are the relevant counterfactuals? Well there are at least three that we can identify as being true that suggest that there is more than one possible future compatible with the past, and hence suggest that the world is indeterministic: (1) we think, say we label a part of the tower "a", that while the region of the tower a is located at space-time point p it could have been located at space-time point r. And (2), we think that while the outside elbow of the collapse is located at point of space p, it is nevertheless true that the outside of the elbow of collapse could have been located at point q. And (3), we think that it is true to say that a, which actually is on the outside of the elbow of the collapse, could have been on the inside of the elbow. In fact, presumably, we think, given the laws and state of the world in the past, that all these counterfactual statements about what might have been are true, and that these counterfactuals characterize ways that the futures could be different given the state of the world in the past. But even if just one of these counterfactual statements were true this would be sufficient to license the claim that there is more than one possible future compatible with the state of the world in the past; this is enough for the claim of indeterminism. The counterfactuals we think are true in CONTINUUM-no-decay-red are these: (1′) we think that if we labeled an alpha particle A_1, while A_1 is located at point p, it could have been located at point q. (2′), we also think that, while the red alpha particle is located at point p, it could have been located at point q. Here, then, we have two true counterfactuals and these are analogues to (1) and (2) above. We also have a true counterfactual analogous to (3) above: We think that while A_1 is red it might not have been red (some other alpha particle could have been red and A_1 would not have been red). And again, we presumably think that given the laws and the state of the world in the past that all three of these counterfactuals are true, and that it is the truth of these counterfactuals that makes it true to say that the there are many different futures compatible with the state of the world in the past. But again, even if just one of these counterfactuals were true this would be sufficient to license the claim that there are many possible futures compatible with the state of the world in the past, and so would license the claim that the world is indeterministic.

So what about CONTINUUM-no-decay? Well, presumably we should say that there are multiple possible futures given the state of the world in the past, and that consequently determinism fails in this world if we think that at least some counterfactuals analogous to the ones we have identified for the tower world and for the CONTINUUM-no-decay-red world are true. The truth of the following statement certainly seems to be compatible with the laws of this world: that A_1, which is actually located (at some time t) at point p, could have been located (at time t) at point q. Likewise, it is true to say that space-time point p, which is actually occupied by A_1, could have been occupied by, say, A_{13}. Thus we have a true counterfactual analogous to (1) and (1′). What we don't have, of course, is a true

counterfactuals analogous to (2), (2′), (3) and (3′). And it is obvious that we are not going to have true counterfactuals analogous to these, for these counterfactuals require that there be objects in the world that play qualitatively different roles, but by hypothesis, in CONTINUUM-no-decay, there are no qualitative differences between the individual alpha particles, or between the individual points of spacetime. But, surely, there are nevertheless differences between the alpha particles in this world: There are continuum many distinct but qualitatively identical alpha particles, and each alpha particle is located in a different place after 13 years. In fact they all have different velocities after 13 years. And the laws do not determine which trajectory any given alpha particle will traverse. Again, the fact that it is undoubtedly true, given the laws of this world and the current state of this world, to say that A_1, which is actually located (at some time t) at point p, could have been located (at time t) at point q, makes it clear that there are many possible futures given the state of the past. Hence, the truth of this counterfactual would seem to be sufficient for a judgment of indeterminism for this world. So it would seem that futures can differ just over which objects are playing which roles, even when those roles are qualitatively the same.[16] And so, again, sophisticated determinism cannot be right.

One might object to the reasoning here not only on the grounds that the examples are far fetched because of the notion of co-location assumed, which I will return to later, but also on the grounds that I have misidentified the counterfactuals that are relevant for determinism. Perhaps the objection would go as follows: it is precisely the fact that we only have a counterfactual analogous to (1) and (1′) in CONTINUUM-no-decay, and none analogous to (2), (2′), (3) and (3′), that makes it reasonable to reject the argument that CONTINUUM-no-decay is indeterministic. Such an objection will have to come with some argument for why the truth of a counterfactual like 1 is not sufficient for a judgment of indeterminism. I take the truth of 1 to justify the claim that there is more than one possible future compatible with the past, and from there it is a quick step to indeterminism given that we agree with the intuitive conception of determinism laid out back in Section 2. So the objection will either have to motivate why the truth of 1 doesn't justify the claim that there is more than one possible future given the past, or it will have to explain why the intuitive conception of determinism is not correct.

I don't think there is a plausible way to mount the objection that the truth of a counterfactual of the kind (1) or (1′) doesn't justify the claim that there is more than one possible future given the past. A given alpha particle has a given trajectory, it is located at a given set of points of space, but we are granting that it is true that it could have had a different one and have been located at a different set of

[16] In fact, shouldn't it be tremendously easy for A_1 to play A_{13}'s role? All they have to do is switch places, the rest will be achieved by "just being themselves". I don't know if we can classify all cases of futures that purportedly differ over which objects are playing which qualitatively identical roles as genuinely different futures. What of PAIR–no-decay (not PAIR, just to bypass the issue of particle creation again)? The question would seem to be whether the relevant modal claims come out true given the laws and the state of the world prior to 13 years. Could A_1, which is actually playing the role of A_1, have played to role of A_2? Well it's a bit harder to get a handle on the answer here than it was in CONTINUUM-no-decay; there at least we could imagine that A_1 was, instead of being located at p, actually located at q—not a qualitative difference, but a difference all the same. But there is nothing like this to point to in PAIR-no-decay. I suppose it comes down to making sense of whether A_1, which actually is A_1 could have been A_2. If this is true, given the laws and state of the world prior to 13 years then I would submit that this world is indeterministic.

points of space. Why is this not a difference and hence a different possible future? Once one grants that we can have distinct but not qualitatively distinct futures (i.e. once we grant that there are haecceitistic differences in futures), which as we have seen, has already been granted by Melia and Skow, and has to be granted by those who want to maintain that the tower world is indeterministic, then what grounds have we for thinking that there are not haecceitistic differences between futures in this example also. Once we buy into haecceitistic differences, it seems to me that we have no grounds for claiming that CONTINUUM-no-decay does not exhibit haecceitistically different futures. And once we grant that, provided we grant the intuitive conception of determinism we get indeterminism in the CONTINUUM-no-decay world.

Should someone mount an objection against the intuitive conception of determinism, and argue that sometimes, even when there is more than one distinct future compatible with the past then determinism sometimes still reigns? Only if one has provided a criterion for when different futures should count as determinism violating futures and when they should not, and one has provided a justification for this criterion. It is not clear to me that this can be done, and I think it is clear that at the very least Melia and Skow haven't provided that justification.

8.3 Sophisticated determinism rejected

I've argued in Section 8.1 that contrary to sophisticated determinism we can have violations of determinism when we have particle creation in duplicate worlds. I have also argued, again contrary to sophisticated determinism, in Section 8.2 that we can have violations of determinism where futures differ over which objects play which roles even when these roles are qualitatively identical. Where does this leave us with respect to the hole argument and the charge of indeterminism there. We should accept that the worlds of the hole argument are indeterministic. In fact the arguments of Section 8.1 are sufficient to establish this: once we have argued that object creation is a red herring we will in general not have the resources to block the charge of indeterminism in the hole argument, and this is independent of whether we accept or reject the argument that we should count futures that differ just in terms of which objects are playing which qualitatively identical role as violating determinism, for in general in the hole argument the futures that differ over which objects are playing which roles will be ones in which those roles are qualitatively different.

9. CONCERNS AND CONCLUSIONS

Are the examples of the previous section just too far fetched to be compelling? CONTINUUM-no-decay and PAIR-red-blue-no-decay both use this rather unattractive feature of objects co-locating in space through time. Perhaps this is so unphysical a feature and so counterintuitive that we should not expect a notion of determinism to give us the "correct" judgments in such cases. After all, I think it would be absurd to argue against a definition of determinism because it gives the

wrong verdict in cases that use trolls and magic. But I don't think these examples are that bad—we do countenance field ontologies in which it is perfectly possible to talk of different fields defined in the same regions of space or space-time. And without too much trouble we could reformulate these examples in terms of fields. We can, for example, imagine that we have two fields defined everywhere in space, that initially have magnitude of zero throughout space at an initial time. We then imagine a law that decrees that such a state can exist for exactly 13 years after which the values of one field will start to increase everywhere it is defined while the values of the other will start to decrease everywhere it is defined.[17] Would such examples be more compelling? Possibly for some, although I am not sure why we should be unwilling to countenance the examples as I present them.

If these examples are simply too far fetched then for now I just have the intuitions of Section 7 to fall back on. These intuitions seem to me to lend strong support to the view that driving the wedge between the hole and the tower using sophisticated determinism is not right. I think we should accept the basic intuitive idea of determinism that maintains that determinism fails when there is more than one physically possible future compatible with the past. Doing so is the safest way to prevent us from giving characterisations of determinism that mask genuine violations. If we do this then I think we should simply bite the bullet in GR and call it what it is: it is an indeterministic theory; the hole argument shows that it is indeterministic. But not all violations of determinism are on a par: some are imperceptible and do not affect our prediction tasks, and some are more worrying. If the cosmic censorship hypothesis fails we will have to recognize that GR is indeterministic in a much more worrying way than we do now. So when is a theory indeterministic? Precisely when there is more than one physically possible future given the present (or past). We get different degrees, senses, or worrying forms of indeterminism depending on just how different these physically possible futures are. Indeterminism arising from classes of worlds that are global duplicates is less worrying than indeterminism that arises from classes of worlds that are not global duplicates. And this is because the alternative futures given the current state in non-global-duplicate violations of determinism are very different, while in global duplicate violations the different futures may be imperceptibly different. Moreover, even within determinism violating classes of duplicate worlds there may be more and less worrying forms of indeterminism. Suppose it turns out that we can drive a well motivated wedge between the tower case and the hole case, would this mean that the position advanced here is no longer defensible? Maybe, but it is not obvious for such a well motivated distinction, if we find one, may give us an argument as to why the varieties of indeterminism that arise from classes of worlds that are global duplicates are not all on a par. And it will further show us why the form of indeterminism in the hole argument is truly unworrying—

[17] One can imagine someone objecting that the example is uncompelling on the grounds that it would be preferable to speak of multiple components of a single field rather than multiple overlapping fields existing in the same region of space. But this is just to provide a different interpretation of your theory—and if you do this you very well may have that on one interpretation of the theory it is deterministic, while on the other it is indeterministic. Still, on the interpretation that commits us to overlapping fields we would have indeterminism, and we should be unsurprised by the fact that whether you think a theory is deterministic or not will be affected by how you interpret your theory.

something we all believe anyway on the grounds that hole diffeomorph worlds are global duplicates.

ACKNOWLEDGEMENT

I would like to thank Gordon Belot and Tom Cuda for helpful comments on earlier drafts.

REFERENCES

Belot, G., 1995. New work for counterpart theorists: determinism. British Journal for the Philosophy of Science 46, 185–195.

Brighouse, C., 1997. Determinism and modality. British Journal for the Philosophy of Science 48, 465–481.

Butterfield, J., 1989. The hole truth. British Journal for the Philosophy of Science 40, 1–28.

Earman, J., Norton, J., 1987. What price spacetime substantivalism? The hole story. British Journal of Philosophy of Science 38, 515–525.

Lewis, D., 1986. On The Plurality of Worlds. Blackwell, New York.

Melia, J., 1999. Holes haecceitism and two conceptions of determinism. British Journal for the Philosophy of Science 50, 639–664.

Skow, B., 2004. The hole argument. Unpublished.

CHAPTER 9

Conventionality of Simultaneity and Reality

Vesselin Petkov*

Abstract

An important epistemological lesson can be learned from the impossibility to determine the one-way velocity of light and the immediate implication that simultaneity is conventional. The vicious circle—to determine whether two distant events are simultaneous we need to know the one-way velocity of light between them, but to determine the one-way velocity of light we need to know that the two events are simultaneous—is an indication of the need for a profound change of our view on reality.

1. INTRODUCTION

After all that has been written on the conventionality in determining the one-way velocity of light and the conventionality of simultaneity (see for instance Poincaré (1898), Einstein (1952), Eddington (1975), Reichenbach (1958), Grünbaum (1973), Janis (SEP), Malament (1977), Sarkar and Stachel (1999), Winnie (1970a, 1970b), Weingard (1972), Petkov (1989), Salmon (1977), Ohanian (2004), Jammer (2006)) one wonders what more can be added to this issue. It turns out, however, that an important aspect has not been sufficiently explored—the link between conventionality of simultaneity and reality. In 1972 Weingard first made this link but only barely mentioned it by devoting a single sentence to it: "But while distant simultaneity is a matter of convention, being real, I take it, cannot be merely a matter of convention" (Weingard, 1972, p. 120). As that link has remained unexplored so far, excluding my own attempts to go a little further (Petkov, 1989, Section 3; Petkov, 2005, Section 5.6.1; Petkov, 2007, Section 2.2), the purpose of this paper is to examine the implications of the issue of conventionality of simultaneity for our view on reality.

* Department of Philosophy, Concordia University, 1455 De Maisonneuve Boulevard West, Montreal, Quebec, Canada H3G 1M8
E-mail: vpetkov@alcor.concordia.ca

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00009-0

© Elsevier BV
All rights reserved

Section 2 analyzes the origin and physical meaning of the vicious circle in attempting to determine the one-way velocity of light and any one-way velocity. Section 3 demonstrates that since the pre-relativistic view on reality is formulated in terms of absolute simultaneity it is directly affected by the relativistic view on simultaneity—that simultaneity of distant events is both conventional and relative.

2. ANY ONE-WAY VELOCITY AND SIMULTANEITY ARE CONVENTIONAL

In 1898 Poincaré first realized that any measurement of the velocity of light is based on an implicit assumption, namely "that light has a constant velocity, and in particular that its velocity is the same in all directions. That is a postulate without which no measurement of this velocity could be attempted. This postulate could never be verified directly by experiment" (Poincaré, 1898, p. 220). Seven years later Einstein arrived at the same conclusion. In the section 'Definition of Simultaneity' of his 1905 paper he discussed the introduction of a common time at two distant points A and B: "We have not defined a common 'time' for A and B, for the latter cannot be defined at all unless we establish *by definition* that the 'time' required by light to travel from A to B equals the 'time' it requires to travel from B to A" (Einstein, 1952, p. 40). It is clear that "by definition" Einstein meant "by convention". Therefore, according to Poincaré and Einstein the magnitude of the one-way velocity of light cannot be discovered by experiment and should be determined by convention.

To see why the one-way velocity of light cannot be determined experimentally assume that we are trying to do just that—to measure the velocity of light from a point A to another point B. To do that we obviously need to know the distance between A and B and the time for which light propagates from A to B. In order to measure that time the clocks at A and B should show the same readings *simultaneously*, i.e. they should be synchronized. But how can that be done? One can use two methods to synchronize the clocks at A and B. The first is to send a light (or any other) signal from A to B whose one-way velocity is *known*. Hence we arrive at a vicious circle—to determine the one-way velocity of light propagating from A to B the clocks at these points should be synchronized, but to synchronize the clocks the one-velocity of light should be known beforehand.

The second method to synchronize the clocks at A and B is the so called slow transport of a third clock C from A to B—the C-clock is initially synchronized with the A-clock and then slowly transported to point B where the B-clock is synchronized with the third clock. It is called "slow transport" to imply that the time dilation that the C-clock undergoes should be neglected. However, neglecting it would mean missing the whole point in the synchronization of distant clocks by a third clock. No matter how small the time dilation might be, if we attempt to calculate it we arrive at the same vicious circle as in the case of the first method: to determine the magnitude of time dilation in order to synchronize the A and B

clocks we should know the one-way velocity of the *C*-clock, but to measure that velocity the *A* and *B* clocks should be synchronized in advance.

One might object that in the case of the second method of synchronizing two distant clocks the vicious circle can be avoided if an observer at rest in clock *C*'s reference frame[1] uses the *C*-clock itself to measure the time of its journey from *A* to *B*, not the clocks at *A* and *B*. Then by knowing the distance between *A* and *B* one can calculate the one-way velocity of *C*. I believe the problem with this objection is obvious—the distance between *A* and *B* is relativistically contracted for the observer in *C*'s reference frame. In order to determine the magnitude of the length contraction the one-way velocity of *C* should be known and we again arrive at the vicious circle.

Two main conclusions can be drawn from here. First, not only the one-way velocity of light but any one-way velocity[2] is a matter of convention since it cannot be directly measured. Second, the conventionality of the one-way velocity of light implies conventionality of simultaneity of distant events as well; this was evident to both Poincaré (1898, p. 222) and Einstein. These conclusions raise difficult questions about their physical meaning. An obvious question is "If *in reality* the velocity of light in one direction has an objective value, how can it depend on human choices and be a matter of definition (convention)?" Obviously, this question is based on the assumption that the concept of velocity, and therefore the one-way velocity of light as well, has a counterpart in the objective world. Questioning this assumption amounts to questioning the fact that objects are in motion, which means that they move with some (definite) velocities with respect to an inertial reference frame. Poincaré does not seem to have been bothered by this type of questions: "So for the velocity of light a value is adopted, such that the astronomic laws compatible with this value may be as simple as possible" (Poincaré, 1898, p. 221). Neither Poincaré nor Einstein appear to have asked the deep questions: "Why can the one-way velocity of light not be measured?" and "Why is simultaneity conventional?" In any case their position on the *physical meaning* of conventionality of the one-way velocity of light and simultaneity is not known.

As we will see below the profound physical meaning of the vicious circle in determining the one-way velocity of light and the resulting conventionality of simultaneity is that the world cannot be three-dimensional. Had Poincaré and Einstein tried to reveal what causes the conventionality of the one-way velocity of light and simultaneity they might have arrived at the idea that reality is a four-dimensional world before Minkowski (1952). This especially applies to Poincaré who wrote in 1906 that the Lorentz transformations are a rotation in a four-dimensional space with time as the forth dimensions (Poincaré, 1906). It seems, however, he regarded that space merely as a mathematical space that does not represent anything real.

[1] Throughout the paper I will follow the widely accepted practice in relativity to regard 'frame of reference' and 'observer' as synonymous."**The word 'observer' is a shorthand way of speaking about the whole collection of recording clocks associated with one free-float frame.** No one real observer could easily do what we ask of the 'ideal observer' in our analysis of relativity. So it is best to think of the observer as a person who goes around reading out the memories of all recording clocks under his control. This is the sophisticated sense in which we hereafter use the phrase 'the observer measures such-and such' " (Taylor and Wheeler, 1992).

[2] 'Velocity' in this paper means what is meant in the context of the conventionality thesis—'three-dimensional velocity', not the 'four-dimensional velocity' of special relativity.

At first sight, the tough questions posed by the conventionality thesis are not immediately obvious. Any one-way velocity is determined by convention but this does not seem to be so surprising. Velocity is a frame-dependent concept which implies that the concept of velocity does not represent something objective; for this reason it is not so unexpected that it does not have an objective (absolute or frame-independent) value.[3] Similarly, simultaneity is conventional but special relativity showed that it is also frame-dependent (relative) and therefore nothing in the objective world corresponds to the concept of absolute simultaneity.

But this is not the whole story. One may argue that the concept of velocity does represent something objective for the following reason. Velocity is relative, but that frame-dependency does not appear to undermine the belief that with respect to a *single* observer the velocity of a particle reflects an objective fact—the *motion* of the particle relative to the observer. So, how can the one-way velocity of a particle be conventional in *one* reference frame? In a given reference frame the one-way velocity of the particle appears to be an objective fact that should not depend on the choice made by an observer in that frame. However, as we will see in Section 3 this argument and this question turn out to be based on our pre-relativistic intuition. In special relativity the concept of velocity (as the measure of the *motion* of a three-dimensional object) is relative (frame-dependent) and, as will be shown in Section 3, does not reflect anything real. And if this concept does not have an objective counterpart we are free to assign a value of our choice to any velocity. So it does follow from the relativity of velocity that velocity is also conventional.

The situation with simultaneity is the same. The frame-dependency of simultaneity demonstrates that no class of absolutely simultaneous events exists. This means that we are indeed free to choose which events to regard as simultaneous since no class of events is objectively or absolutely privileged. Therefore relativity of simultaneity implies conventionality of simultaneity. The opposite is also true—conventionality of simultaneity implies relativity of simultaneity. As distant simultaneity is conventional it follows that no class of events is objectively privileged as being simultaneous; if such a class of privileged simultaneous events existed, then simultaneity could not be conventional. But as no class of events is absolutely (objectively) simultaneous, different observers in relative motion are not forced (due to the lack of a class of objectively privileged simultaneous events) to share the same class of simultaneous events, which means that simultaneity is not absolute and is therefore relative.[4] I think the fact that relativity of simultaneity implies conventionality of simultaneity and vice versa should be specifically emphasized since any claim that simultaneity is relative but not conventional amounts to a contradiction in terms: there is no objectively privileged class of simultaneous events (due to relativity of simultaneity), but there is an objectively privileged class of simultaneous events (due to the non-conventionality of simultaneity).

[3] There would be a problem with the conventionality thesis only if the value of a frame-independent physical quantity would be a matter of convention.

[4] If two observers in relative motion choose either the standard (Einstein) convention $\varepsilon = \frac{1}{2}$ or the *same* non-standard convention $\varepsilon \neq \frac{1}{2}$ they will have different classes of simultaneous events, which means that simultaneity will be relative for them. There will be no relativity of simultaneity only in one special case when they choose the same class of events to be simultaneous for each of them; this choice corresponds to different values of ε for each of the observers.

The direct link between relativity of simultaneity and conventionality of simultaneity follows from the fact that "in special relativity, the causal structure of space-time defines a notion of a 'light cone' of an event, but does not define a notion of simultaneity" (Wald, 2006). No class of events lying outside of the light cone at an event P can be defined as simultaneous on the basis of causal relations in spacetime, which means that no class of events is objectively privileged or objectively distinct from the other events in that region. For comparison consider the light cone at P. Its three regions—past, future, and outside—are, in terms of causal relations,[5] objectively distinct from one another; that is why the light cone is a frame-independent concept. However, in relativity the concept of simultaneity is still used—for example, length contraction and time dilation cannot be formulated if that concept is not employed. As a result of our insistence on using the concept of simultaneity where it does not reflect anything objective we arrive at the conclusions that simultaneity is relative when two observers in relative motion are considered and that simultaneity is conventional if just one observer is involved. Both conclusions follow from the fact that the status of all events in the area outside of the light cone is the *same*. That is why if simultaneity is relative it is also conventional and vice versa—both relativity and conventionality of simultaneity imply that there is no objectively privileged class of events among all events in the area lying outside of the light cone.[6]

So the conventionality of the one-way velocity of light or the one-way velocity of any particle and of simultaneity is inescapable in the framework of relativity.[7] Despite this, however, the tough questions mentioned above remain. It is sufficient to point out the conventionality of the one-way velocity of light. Unlike the velocities of particles, the velocity of light is absolute (frame-independent). This suggests that the concept of velocity of light reflects something objective. Then, how can the one-way velocity of light be a matter of convention? As we will see in the next section, an even more difficult question is raised when it is taken into account that conventionality of simultaneity would imply conventionality with respect to what exists if reality were a three-dimensional world. The link between conventionality of simultaneity and reality will be explored in the next section which will allow us to arrive at a view on reality that is fully consistent with the

[5] All events in the past light cone can influence the event at P. The event at P can affect all events in the future light cone. No event in the spacetime region lying outside of the light cone can influence the event at P and vice versa.

[6] In this connection it is worthwhile to point out that Malament's theorem (Malament, 1977) merely proves that the standard synchronization $\varepsilon = \frac{1}{2}$ is the simplest one. His result did not disprove the conventionality thesis. If we assume it did, it would follow that the class of simultaneous events (determined by the choice $\varepsilon = \frac{1}{2}$) would be objectively privileged (no conventionality!) and therefore observers in relative motion would share the same class of objectively privileged evens. Hence simultaneity would turn out to be absolute, if distant simultaneity were not conventional. That is why it is indeed a contradiction in terms to say that simultaneity is not conventional, but is relative.

[7] It is natural to ask whether conventionality of any one-way velocity and simultaneity is a feature of only the theory of relativity. In Newtonian physics the first method of synchronizing distant clocks through light (or any other) signals also leads to a vicious circle. However, the second method (slow transport of a third clock) does not lead to such problems since there is no time dilation and length contraction in Newtonian physics. Hence the one-way velocity of light and simultaneity of distant events are not conventional in classical (pre-relativistic) physics. This conclusion also follows from the fact that simultaneity is absolute in Newtonian physics. At any moment of the absolute time there is one class of absolutely simultaneous events that is shared by all observers in relative motion. Due to the privileged status of this class the observers are not free to choose different classes of simultaneous events. By the same argument every single observer is not free to choose which events to be regarded as simultaneous, which means that simultaneity is not conventional. As we will see in the next section what makes the only class of absolutely simultaneous events privileged is the fact that according to the pre-relativistic world view what exists at the present moment is namely the class of absolutely simultaneous events.

conventionality thesis and that provides natural answers to all difficult questions raised by the conventionality thesis.

3. SIMULTANEITY AND REALITY

The real challenge of the conventionality thesis is fully manifested when the issue of what is real is explicitly addressed. Since the time of Aristotle reality has been regarded as a three-dimensional world (Aristotle, 1993). At that time the concept of reality could have been formulated only in terms of what is directly perceived—the observable world. Therefore, until the seventeenth century the three-dimensional world could have been defined as 'everything that we *see* (or can in principle see) *simultaneously* at the present moment'. However, after Rømer determined in 1675 that the velocity of light was finite it became clear that what we see is all past. Then the second view on reality could have defined the three-dimensional world as 'everything that *exists simultaneously* at the present moment'. This second view on reality, called presentism, was fully consistent with the pre-relativistic physics, but is incompatible with relativity.

I believe the reason is obvious—the pre-relativistic view on reality is defined in terms of absolute simultaneity, but according to special relativity simultaneity is both relative and conventional which means that no class of simultaneous events, that can be identified with the three-dimensional world, is objectively privileged. If reality were a three-dimensional world, i.e. a single class of simultaneous events, conventionality of simultaneity would imply that what exists is also a matter of convention which is clearly unacceptable.[8] Therefore the message of the vicious circle involved in any attempt to determine any one-way velocity and simultaneity of distant events is truly profound—reality is not a three-dimensional world. The same conclusion follows from relativity of simultaneity—as two observers in relative motion have different classes of simultaneous events it follows that they have different three-dimensional worlds, which is possible only if these worlds are three-dimensional cross-sections of a real four-dimensional world represented by Minkowski spacetime. Here I would like to emphasize again the link between relativity of simultaneity and conventionality of simultaneity—both relativity of simultaneity and conventionality of simultaneity imply that reality is a four-dimensional world; neither relativity of simultaneity nor conventionality of simultaneity are possible in a three-dimensional world.

To see this and also why the pre-relativistic world view is in unsurmountable contradiction with special relativity, assume that reality were indeed a three-dimensional world. Then it would follow that

[8] On the pre-relativistic (presentist) view everything that exists, exists at the present moment. Therefore it is natural to consider absolute simultaneity at the moment 'now' in order to understand more clearly its physical meaning and why simultaneity is not conventional in the pre-relativistic physics. When it is taken into account that the concept of absolute simultaneity in Newtonian physics implies a single three-dimensional world (defined as everything that exists simultaneously at the present moment) it becomes clear that it is the existence of the simultaneous events at the present moment that makes them absolutely simultaneous or objectively privileged: according to the pre-relativistic world view reality is a single three-dimensional world (the present); therefore there exists a single class of simultaneous events at the moment 'now', which as the only class of simultaneous events is absolute in a sense that it is common to all observers in relative motion. As on the presentist view reality is a single class of absolutely simultaneous events (at the moment 'now'), conventionality of simultaneity would mean that we would be free to determine what is real.

- No relativity of simultaneity would be possible since all observers in relative motion would share the same three-dimensional world and therefore would have the same class of simultaneous events.
- No conventionality of simultaneity would be possible since the three-dimensional world would be the only thing that exists and therefore would be objectively privileged and not a matter of convention.

It is explicit that the conclusion 'neither relativity of simultaneity nor conventionality of simultaneity are possible in a three-dimensional world' follows from the definition of a three-dimensional world—the class of events that are absolutely[9] *simultaneous* at the present moment. In such a desperate situation when the widely accepted presentist view is strongly challenged by special relativity, it appears natural to ask "Is it possible to define the three-dimensional world in such a way that simultaneity is not involved?" The answer is "No". On the presentist view it is only the three-dimensional world that exists at the present moment. If it is not defined as 'all events that are simultaneous at the present moment', the only option for another definition is that the three-dimensional world contains events occurring at different moments of time. This is clearly impossible in the framework of the presentist view since such a three-dimensional world would contain past, present, and future events, not only the existing present events.

Another possibility to challenge the conclusion that special relativity is impossible in a three-dimensional world is to point out that the 1905 formulation of relativity given by Einstein was in terms of the ordinary three-dimensionlist (presentist) view. It is true that special relativity can be equally formulated in the usual three-dimensional language as well as in the four-dimensional language of Minkowski spacetime. However, while both representations of relativity correctly describe the relativistic phenomena, they are obviously not equivalent in terms of the dimensionality of the world. Therefore only one of them adequately represents the world's dimensionality. So the original formulation of special relativity was in three-dimensional language, but it does not mean that it was possible in a three-dimensional world. To see this let us ask what the *physical meaning* of the kinematical relativistic effects is. It becomes immediately evident that they are impossible in a three-dimensional world. Take as an example relativity of simultaneity and assume that reality were a three-dimensional world (the present). As such a world is defined as the class of events that are simultaneous at the present moment it follows that all observers in relative motion would share the same class of simultaneous events since only this class of events would exist at the moment 'now'. This means that simultaneity would be absolute in contradic-

[9] Strictly speaking, it is not even necessary to add 'absolutely' to the definition of a three-dimensional world for two reasons:

- On the presentist view there exists a single class of simultaneous events at the present moment, which as the only class of simultaneous events are absolute because they are common to all observers in relative motion.
- When a three-dimensional world is regarded as a sub-space in Minkowski spacetime one cannot use absolute simultaneity. In this case a three-dimensional world (a three-dimensional cross-section of Minkowski spacetime) is defined merely as 'the class of events that are *simultaneous* at a given moment of an observer's time'. As we saw above due to relativity of simultaneity two observers in relative motion have different three-dimensional worlds which are simply different three-dimensional cross-sections of Minkowski spacetime.

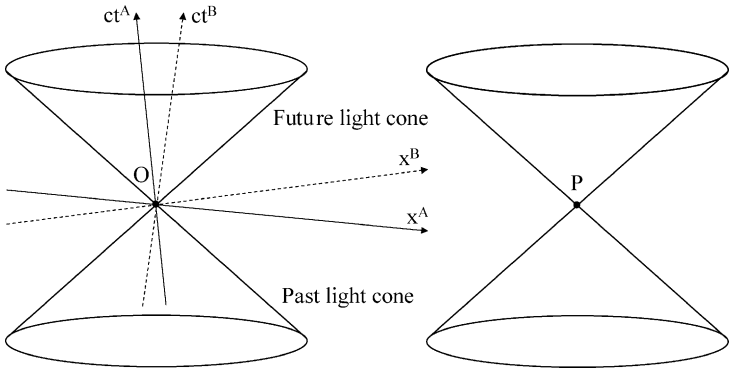

FIGURE 9.1 Two observers A and B in relative motion, who meet at event O, are represented by their worldlines. If it is assumed that what is real is represented by the area outside the light cone at O, it follows that the observers will have the freedom to choose different classes of simultaneous events (represented by their x-axes). The events in the past and future light cone also turn out to be real since they fall in the area lying outside a second light cone at event P which is space-likely separated from event O.

tion with special relativity. Therefore special relativity is indeed impossible in a three-dimensional world.

As the causal structure of spacetime defines a notion of a light cone, not a notion of simultaneity, the most rigorous approach to the issue of reality in the framework of relativity is to ask what is real in terms of the light cone. What is immediately clear is that reality cannot be a three-dimensional world since it is defined in terms of simultaneity—as everything that exists simultaneously at the moment 'now' of an observer's time—whereas the spacetime causal structure does not define such a concept. As a first attempt one can identify the spacetime area lying outside the light cone at event O (Figure 9.1). This choice is dictated by both relativity of simultaneity and conventionality of simultaneity—the area outside of the light cone at O must exist in order that (i) two observers A and B in relative motion could have different instantaneous three-dimensional spaces, i.e. different classes of simultaneous events (which are represented by the x-axes of the observers in Figure 9.1), and (ii) each of the observers could choose (by convention) his own instantaneous three-dimensional space, i.e. his class of simultaneous events from that spacetime area. Put another way, the area outside the light cone at O must exist in order that relativity of simultaneity and conventionality of simultaneity be possible.

The next step would be to ask: what is the status of the events in the past and future light cone? As Weingard (1972) demonstrated, if reality is represented by the area lying outside the light cone at O, by the same criterion events in the past and future light cone are also real since they lie in the area outside a second light cone at P (Figure 9.1). So it follows that all spacetime events are real. This conclusion is inevitable when one asks what the impact of each of the relativistic changes of the concept of simultaneity—relativity of simultaneity and conventionality of simultaneity—on the view on reality is. The relativistic view, according to which

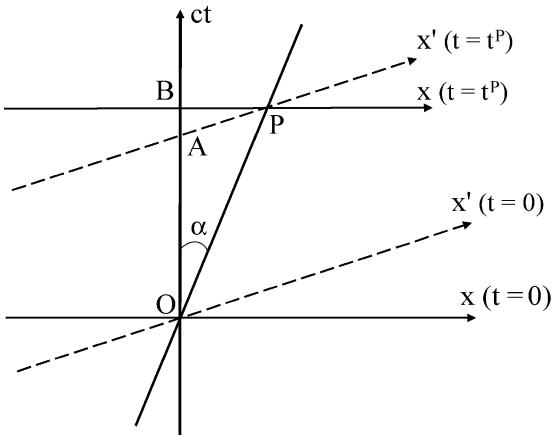

FIGURE 9.2 The worldlines of an observer and a body form an angle α. The time axis is chosen along the worldline of the observer. In a three-dimensional language the two worldlines can be interpreted to mean that the observer and the body are in relative motion. Although their relative velocity can be expressed in terms of the angle α, it does not mean that the relative velocity is completely determined by α. The observer is free to determine by convention whether his instantaneous three-dimensional space (depicted here only by the x axis) is orthogonal to his worldline (x) or not (x'). The relative velocity between the observer and the body depends on this choice.

reality is represented by the four-dimensional Minkowski world, provides complete answers to the difficult questions raised by the conventionality thesis.

In Minkowski world the whole history of every signal or every body is entirely realized in the signal's worldline or the body's worldtube. There are no three-dimensional objects in spacetime and no motion of such objects. That is why the concept of velocity does not have an ontological counterpart. For this reason we are indeed free to choose the value of velocity when we describe Minkowski spacetime in terms of our three-dimensional language. When the motion of a body is described with respect to a given observer what corresponds to the body's velocity is the angle α between the worldlines of the observer and the body (Figure 9.2). However, α is not the ontological counterpart of the relative velocity since the relative velocity depends not only on α, but also on whether the instantaneous three-dimensional space of the observer is chosen to be orthogonal to the observer's worldline or not. That is why the velocity of the body with respect to the observer is a matter of convention. This is clearly seen in Figure 9.2. Imagine that the observer decides to measure the velocity of the body, which travels from event O to event P. Depending on whether or not the observer's instantaneous three-dimensional space is orthogonal to his worldline, the time it takes the body to reach event P will be either OB or OA. Therefore, the body's velocity is indeed determined by convention.

The velocity of a light signal is conventional for the same reason—the observer is free to choose his instantaneous three-dimensional space to be either orthogonal to his worldline or to form an angle with it. The fact that the velocity of light is not

frame-dependent is a result of the frame-independency of the concept of a light cone but that does not affect the conventionality of the light velocity due to the observer's freedom to chose the angle between his worldline and his instantaneous three-dimensional space.

CONCLUSION

The epistemological lesson that can be learned from the impossibility to determine the one-way velocity of light and the immediate implication that simultaneity is conventional demonstrates that every time when we arrive at a vicious circle some of our views should be drastically changed. And indeed the fact that the one-way velocity of light and simultaneity of distant events are conventional has turned out to have a profound meaning—reality is a four-dimensional world represented by Minkowski spacetime. There are no moving light signals or three-dimensional bodies in this four-dimensional world and when we describe it in our three-dimensional language in terms of motions, the velocities of these signals and bodies are determined by convention since they do no represent anything real.

ACKNOWLEDGEMENTS

This chapter reflects the talk I gave at the Second International Conference on the Ontology of Spacetime which was held at Concordia University, Montreal, on June 9–11, 2006. I would like to thank all colleagues who attended the talk and took part in the discussion after it. I would also like to acknowledge helpful suggestions and constructive criticism by an anonymous referee.

REFERENCES

Aristotle, 1993. On the Heavens, Book I. In: Adler, M.J. (Ed.), Great Books of the Western World, vol. 7. Encyclopedia Britannica, Chicago, pp. 357–405.

Eddington, A., 1975. The Mathematical Theory of Relativity, 3rd ed. Chelsea, New York, pp. 15–16.

Einstein, A., 1952. On the electrodynamics of moving bodies. In: Lorentz, H.A., Einstein, A., Minkowski, H., Weyl, H. (Eds.), The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity. Dover, New York, pp. 37–65.

Grünbaum, A., 1973. Philosophical Problems of Space and Time, 2nd ed. D. Reidel, Dordrecht, Boston.

Jammer, M., 2006. Concepts of Simultaneity: From Antiquity to Einstein and Beyond. Johns Hopkins University Press, Baltimore.

Janis, A., SEP. Conventionality of simultaneity. Online Stanford Encyclopedia of Philosophy (and the references therein).

Malament, D., 1977. Causal theories of time and the conventionality of simultaniety. Noûs 11, 293–300.

Minkowski, H., 1952. Space and time. In: Lorentz, H.A., Einstein, A., Minkowski, H., Weyl, H. (Eds.), The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity. Dover, New York, pp. 75–91.

Ohanian, H., 2004. The role of dynamics in the synchronization problem. American Journal of Physics 72, 141–148.

Petkov, V., 1989. Simultaneity, conventionality and existence. British Journal for the Philosophy of Science 40, 69–76.
Petkov, V., 2005. Relativity and the Nature of Spacetime. Springer, Berlin, Heidelberg, New York.
Petkov, V., 2007. Relativity, dimensionality, and existence. In: Petkov, V. (Ed.), Relativity and the Dimensionality of the World. Springer, Berlin, Heidelberg, New York.
Poincaré, H., 1898. La mesure du temps. Revue de métaphysique et de morale 6, 1–13. English translation in: The Value of Science, The Modern Library, New York, 2001, pp. 210–222.
Poincaré, H., 1906. Sur la dynamique de lélectron. Rendiconti del Circolo Matematico di Palermo 21. Reprinted in H. Poincaré: La mécanique nouvelle, Éd. Guillaume (Ed.), Gauthier-Villars, Paris, 1989, pp. 18–76.
Reichenbach, H., 1958. The Philosophy of Space and Time. Dover, New York.
Salmon, W., 1977. The philosophical significance of the one-way speed of light. Noûs 11, 253–292.
Sarkar, S., Stachel, J., 1999. Did Malament prove the non-conventionality of simultaneity in the special theory of relativity? Philosophy of Science 66, 208–220.
Taylor, E.F., Wheeler, J.A., 1992. Spacetime Physics, 2nd edn. W.H. Freeman, New York, p. 39.
Wald, R.M., 2006. Resource letter TMGR-1: Teaching the mathematics of general relativity. American Journal of Physics 74, 471–477.
Weingard, R., 1972. Relativity and the reality of past and future events. British Journal for the Philosophy of Science 23, 119–121.
Winnie, J., 1970a. Special relativity without one-way velocity assumptions: Part I. Philosophy of Science 37, 81–99.
Winnie, J., 1970b. Special relativity without one-way velocity assumptions: Part II. Philosophy of Science 37, 223–238.

CHAPTER 10

Pruning Some Branches from "Branching Spacetimes"

John Earman*

Abstract

Discussions of branching time and branching spacetime have become common in the philosophical literature. If properly understood, these conceptions can be harmless. But they are sometimes used in the service of debatable and even downright pernicious doctrines. The purpose of this chapter is to identify the pernicious branching and prune it back.

1. INTRODUCTION

Talk of "branching time" and "branching spacetime" is wide spread in the philosophical literature. Such expressions, if properly understood, can be innocuous. But they are sometimes used in the service of debatable and even downright pernicious doctrines. The purpose of this paper is to identify the pernicious branching and prune it back.

Section 2 distinguishes three types of spacetime branching: individual branching, ensemble branching, and Belnap branching. Individual branching, as the name indicates, involves a branching structure in individual spacetime models. It is argued that such branching is neither necessary nor sufficient for indeterminism, which is explicated in terms of the branching in the ensemble of spacetime models satisfying the laws of physics. Belnap branching refers to the sort of branching used by the Belnap school of branching spacetimes. An attempt is made to situate this sort of branching with respect to ensemble branching and individual branching. Section 3 is a sustained critique of various ways of trying to implement individual branching for relativistic spacetimes. Conclusions are given in Section 4.

* Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, USA

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00010-7

© Elsevier BV
All rights reserved

2. THREE TYPES OF SPACETIME BRANCHING

The issue of futuristic (respectively, historical) determinism is sometimes posed in terms of the question of whether the laws of physics allow branching in the future (respectively, the past).[1] The relevant sense of "branching" is what I will call *ensemble branching*. The relevant ensemble is the collection of spacetime models (worlds, histories, ...) satisfying the laws.[2] Future branching means that the ensemble contains models that are isomorphic at a time (or isomorphic over some finite stretch of time, or isomorphic for all times less than or equal to a given time) but nonisomorphic for future times;[3] past branching is understood analogously. The isomorphism at issue may be construed either as literal identity or as a counterpart relation.[4] This distinction will be relevant below, but it can be ignored for the nonce. The presence of branching in the ensemble of spacetime models satisfying given physical laws is a *prima facie* indication that these laws are not deterministic. The qualification is needed because although the branching creates a presumption of the failure of determinism, the presumption can be rebutted by identifying gauge freedom; two examples help to illustrate the point.

The first example comes from the context of the special theory of relativity (STR). Consider the source-free Maxwell equations for electromagnetism written in terms of the electromagnetic (four-) potentials.[5] The instantaneous values of the potentials and their first time derivatives do not suffice to fix a unique solution. This failure of non-uniqueness is not taken as an insult to Laplacian determinism because (according to the standard story) the electromagnetic potentials are mere auxiliary devices. The genuine physical magnitudes are electric and magnetic field strengths, and formulated in terms of these quantities the Maxwell equations do admit a well-posed initial value problem, securing Laplacian determinism for worlds described purely in terms of (source free) electromagnetic fields. The values of the electromagnetic potential correspond many-one to the values of the electric and magnetic field strength, and the transformations among the many that correspond to the one are gauge transformations. In this example all of the spacetime models use a single fixed spacetime background, Minkowski spacetime. The second example moves from the context of STR to the context of the general relativity theory (GTR) where the spacetime structure can vary from model to model.[6] The initial value problem for the source-free Einstein gravitational field equations admits a solution that is guaranteed to be unique (locally

[1] For a recent overview of the fortunes of determinism in modern physics, see Earman (2007a).

[2] Since for present purposes it makes little or no difference whether one works with models, or worlds, or histories, I will feel free to shift back and forth amongst them.

[3] The first type of branching is relevant to the issue of Laplacian while the others are relevant to weakened cousins of Laplacian determinism.

[4] For those familiar with the language of differential geometry, think of a spacetime model as having the form $\langle \mathcal{M}, O_1, O_2, \ldots, O_N \rangle$ where $\mathcal{M}$ is a differentiable manifold and the O_i are geometric object fields on $\mathcal{M}$ that characterize either the structure of spacetime or the matter-fields that inhabit spacetime. The models for Newtonian, special relativistic, and general relativistic theories can all be put in this form. Two such models $\langle \mathcal{M}, O_1, O_2, \ldots, O_N \rangle$ and $\langle \mathcal{M}', O'_1, O'_2, \ldots, O'_N \rangle$ are isomorphic in the relevant sense iff there is a diffeomorphism $d : \mathcal{M} \to \mathcal{M}'$ such that $d^* O_i = O'_i$ for all i, where $d^* O_i$ denotes the "drag along" of O_i by d.

[5] The spacetime models have the form $\langle \mathbb{R}^4, \eta_{ab}, \phi_1, \phi_2, \phi_3, \phi_4 \rangle$ where η_{ab} is the Minkowski metric and the ϕ_i are scalar fields on $\mathbb{R}^4$.

[6] The spacetime models have the form $\langle \mathcal{M}, g_{ab} \rangle$ where g_{ab} is a Lorentz signature metric on $\mathcal{M}$. See Section 3.2 below.

in time) only up to a diffeomorphism.[7] The "up to" qualification is not taken to signal a failure of Laplacian determinism because (according to one persuasive interpretational stance) the diffeomorphism invariance of the theory is to be interpreted as a gauge symmetry. From here on I will assume that gauge freedom has been removed so that indeterminism can be safely inferred from ensemble branching.[8]

Note that ensemble branching can obtain even though no spacetime model in the ensemble has a structure that, by any reasonable standard, can be deemed to be branching. But it is precisely the branching in individual spacetime models that seems to be presupposed by some versions of the many worlds interpretation of quantum mechanics (QM). Individual spacetime branching is explicitly advocated by Storrs McCall (1994, 1995, 2000) and Roy Douglas (1995). And it is considered (but rejected) by Roger Penrose (1979). This nonensemble *individual branching* will be the main target of my pruning operation.

But before I set my sights on this target, I want to introduce a third sense of branching—what I will call *Belnap branching*—that lies somewhere between ensemble branching and individual branching, although its exact location is not easy to pin down. The Belnap school's endorsement of branching begins with an anti-Lewisian stance.[9] In *On the Plurality of Worlds* David Lewis informs the reader that he rejects what he dubs 'branching' of possible worlds but accepts 'divergence'. The difference between the two is characterized as follows.

> In branching, worlds are like Siamese twins. There is one initial spatiotemporal segment; it is continued by two different futures—different both numerically and qualitatively—and so there are two overlapping worlds. One world consists of the initial segment plus one of its futures; the other world consists of the identical initial segment plus the other future.
> In divergence on the other hand, there is no overlap. Two worlds have duplicate initial segments, not one that they share in common. I, and the world I am part of, have only one future. There are other worlds that diverge from us. … Not I, but only some very good counterparts of me, inhabit these other worlds. (Lewis, 1986, p. 206)

Contra Lewis, the Belnapians insist that the openness of the future requires 'branching' and not mere 'divergence'. This insistence might be interpreted as an endorsement of the branching of individual spacetimes, but a more conservative reading is possible. On this reading, Lewis' 'divergence' and 'branching' are both subsumed under ensemble branching. In endorsing the former and rejecting the latter Lewis is asserting that in ensemble branching the relation of isomorphism should not be interpreted as identity but as a counterpart relation. The Belnap school is read as insisting that indeterminism requires that the relation of

[7] See Section 3.4 below for more details.

[8] Note that on this interpretational stance neither the spacetime metric nor any of the scalar invariants formed from the metric count as observables or gauge invariant quantities. The issue of how to characterize the gauge-independent content of GTR is controversial; see Earman (2006).

[9] The Belnap school includes, of course, Nuel Belnap, his collaborators (Mitchell Green, Michael Perloff, and Ming Xu), and those who accept his framework for treating branching time and branching spacetime (e.g., Thomas Müller and Thomas Placek). Representative references from this school include Belnap (1992, 2002, 2003), Belnap et al. (2001), Kowalski and Placek (1999), Müller (2004), Placek (2000) and Placek and Müller (2005).

isomorphism must be interpreted as identity and, thus, as opting for 'branching' over 'divergence'. Leaving aside the merits of this dispute, the important point for present purposes is that this dispute does not require taking a stance on branching of individual spacetimes.

In support of this conservative reading, there are passages in which the Belnapians seem to be rejecting individual branching in favor of ensemble branching. Thus, in Belnap et al.'s *Facing the Future* we read:

> [A]lthough we use the phrase 'branching time' because of its fixed place in the literature, we never, ever mean to suggest that time itself—which is presumably best thought of as linear—ever, ever, 'branches.' The less misleading phrase ... is 'branching histories,' with an essential plural to convey that it is the entire assemblage of histories that has a branching structure. (Belnap et al., 2001, p. 29, fn 1)

And what holds for branching times presumably holds *inter alia* for branching spacetimes. Note, however, that in the very same passage Belnap et al. refer the reader to McCall (1994) for "similar ideas" on branching spacetimes. This reference is explicit in endorsing individual branching, and it also returns the complement by referring to Belnap (1992) as a source of inspiration (see McCall, 1994, p. 4, fn 4). The reader is left wondering whether Belnap branching spacetimes involve only ensemble branching or something more.

Another characteristic feature of the Belnap school—the denial of the "thin red line"—strongly hints of something beyond ensemble branching. The doctrine of the thin red line is the doctrine that, as of an indeterministic moment ("branch point"), exactly one future reaching branch is the actual future. John MacFarlane has opined that "positing a thin red line amounts to giving up on objective indeterminism" (2003, p. 325). This is a non-sequitur if indeterminism is explicated as ensemble branching, which I claim, is the sense needed to illuminate issues of determinism and indeterminism in physics. On this explication futuristic indeterminism means no more and no less than that there is more than one possible future history compatible with the combination of the laws of physics and the history up-to-now.[10] It certainly does not follow that (as of now) there is no fact of the matter as to which of the possible futures is the actual future. To get that result some additional piece of metaphysics would have to be added. It might be supplied by Storrs McCall's model of indeterminism in terms of individual spacetime branching and the attrition of branches with the advance of "now" (see Section 3.1 below). An alternative piece of metaphysics that would bridge the non-sequitur without involving a commitment to individual branching would be supplied by C.D. Broad's (1923) growing block universe picture of spacetime according to which the block is built up as successive layers of "now" come into existence. On this picture there is no line, thin or thick, red or any other color, that marks the course of the future since there literally is no future. But note well that this is so regardless of whether or not the growth of Broad's growing block is indeterministic; this severing of the

[10] I do not understand what MacFarlane's qualifier "objective" is supposed to add, other than to emphasize that the indeterminism is not purely epistemic. But this is already captured by ensemble branching when the ensemble is taken to be the models of the physical laws.

link between indeterminism and the "openness" of the future is contrary to the spirit of the Belnap school.

Perhaps the no-thin-red-line doctrine is not supposed to be given an ontological reading but only a semantic reading on which some future tensed statements are neither true nor false. Consider the semantic rule (*R*): as of an indeterministic moment, a statement asserting the future occurrence of an event of type *E* (e.g. 'sea-battle') is true (respectively, false, indeterminate) iff a token *e* of *E* is present in every (respectively, no, some but not every) future reaching branch. Branching in individual spacetimes is not needed to get the consequence that, as uttered now, "There will be a sea battle" has an indeterminate truth value. What is needed is (i) futuristic indeterminism with respect to sea battles, i.e. ensemble branching of physically possible future histories, with some branches containing sea battles and some containing nonesuch, and (ii) the decision to take 'branch' in (*R*) to range over possible future histories compatible with the combination of the laws of physics and the history up-to-now.

But although the semantic reading of the no-thin-red-line doctrine does not necessarily imply a commitment to individual branching, it does not ring true without some additional ontological underpinning. Suppose that Broad's growing block universe is rejected—as I think it should be since it is difficult if not impossible to reconcile it with the relativistic conception of spacetime—and suppose further that branching in individual spacetimes is eschewed. Then the denial of truth values to future contingents has a hollow ring to it. For on the stated suppositions there are no actual future reaching branches sprouting from the trunk of the actual world up to now; there is only a single actual trunk extending into the future—in this sense not only is there a thin red line but also a broad red band painting the unbranching trunk of the actual world. And a natural rule for assigning truth values to future tensed statements is (*R*) with 'branch' taken to range over actual branches. This makes "There will be a sea battle" true in case the (hypothesized) unique actual future contains a sea-battle token and false otherwise. Of course, it is consistent to adopt the form of (*R*) that takes 'branch' to range over possible future histories compatible with the combination of the laws of physics and the history up-to-now. But if indeterminism *per se*, without the aid of an evolving block universe or individual branching, supports the adoption of this form of (*R*), then it equally supports the analogous version of (*R*) for past tensed statements, yielding a version of presentism on which contingent statements about the past as well as the future lack truth values. But the Belnap school has no truck with presentism; indeed, it posits that while there is branching in the future, there is no branching in the past. This posit does not follow from indeterminism if indeterminism is explicated as ensemble branching and the indeterminism is the indeterminism of time reversal invariant laws. For laws that are time reversal invariant, futuristic and historical determinism stand or fall together; in branching jargon, either there is ensemble branching in the ensemble of physically possible histories in both the past and the future or in neither. Time reversal invariance is a property of all of the known fundamental laws of physics, save for some that govern the weak interactions of elementary particles, processes that are presum-

ably irrelevant to explaining any of the temporal asymmetry we observe at the macrolevel.

In sum, both the ontological and the semantic readings of the no-thin-red-line doctrine seem to presuppose something beyond indeterminism as explicated by ensemble branching.[11] This is implicitly acknowledged in *Facing the Future* where it is announced that "indeterminism" is to be understood in a sense that validates the no-thin-red-line doctrine, a future-branching tree structure for time and spacetime, and the other postulates imposed on this structure (see Belnap et al., 2001, pp. 133–141). While there is nothing objectionable *per se* to such a procedure, it leaves unclear what the connection is between the sense of indeterminism needed to satisfy these demands and the sense of indeterminism that (I claim) is relevant to understanding physics.[12]

Since I have been unable to get a fix on what Belnap branching involves, all I can say for the present is this: insofar as Belnap branching eschews individual branching, then for present purposes I have no quarrel with it; but insofar as it is committed to individual branching, it will be subject to the pruning operation to be undertaken in the next section.

3. PROBLEMS WITH INDIVIDUAL BRANCHING

3.1 Faulty motivation

McCall (1994, 1995, 2000) takes individual branching as a way of modeling indeterminism. I claim that there is no necessary connection, in either direction, between determinism and individual branching. As noted above, ensemble branching can obtain even when each and every model (world, history) in the ensemble describes an unbranching spacetime. Thus, if indeterminism is adequately explicated by ensemble branching, indeterminism does not entail individual branching. A detailed examination of determinism and indeterminism in modern physics (see Earman, 2006, 2007a, 2007b) supports the antecedent.

In the other direction, branching in individual spacetimes by itself need not entail indeterminism. If there is no branching in the ensemble of models satisfying given laws, then these laws are deterministic. But it is *a priori* possible that the ensemble contains models whose individual spacetimes branch; that is, although

[11] MacFarlane's claim that "positing a thin red line amounts to giving up on objective indeterminism" is a not-so-subtle attempt to put his opponents on the defense by attributing to them a "posit" of a metaphysical entity—the thin red line. On the contrary, it is those deniers of the thin red line—like MacFarlane and the Belnap school—who hold that future tensed but not past tensed statements may lack a determinate truth value that need to rely on some posit beyond indeterminism in the ensemble branching sense to motivate their semantic rules. MacFarlane's own account of future contingents is much more intricate than anything considered here; in particular, he proposes to relativize truth assignments to "contexts of assessments."

[12] The postulates used by the Belnap program seem to place *a priori* constraints on physics. For example, members of the Belnap school speak of "choice points," which are to be thought of as the loci of indeterministic influences that radiate outwards along or within the future null cones of these loci. On this way of conceptualizing indeterminism, the members of the ensemble of physically possible models will typically agree at all points except those that lie on or inside the future null cones of the choice points. I know of no theory in modern physics which will produce this kind of indeterminism. Relativistic field theories, whether classical or quantum, typically entail lawlike connections among relatively spacelike events.

the laws plus initial conditions pick out a unique temporal evolution, that evolution might involve a spacetime that literally branches. (In fact, however, one means of implementing individual branching in general relativistic spacetimes does lead to indeterminism; but the problem is that it leads to too much indeterminism (see Section 3.4 below).) Indeterminism can emerge from branching in individual spacetimes by adding a chance or random falling off or withering away of branches as "now" creeps up the trunk of the branching spacetime tree.[13] But in addition to the use of the dubious metaphysical notion of a shifting "now," this approach makes the analysis of indeterminism hostage to the even less well understood concepts of chance and randomness.

Finally, it will be seen in Section 3.4 below that the most plausible way of implementing branching for relativistic spacetimes is incapable of modeling some forms of indeterminism.

3.2 No-go results on topology change in classical GTR[14]

Mathematical details aside, the intuitive idea of a future branching spacetime seems clear enough: spacetime is like a tree in that it starts with a main trunk which sends off future reaching branches, each of which in turn may send off its own branches, etc.

As intuitively appealing as the idea of literally branching spacetime might seem to be on first impression, there are serious difficulties in implementing it in terms of relativistic spacetime structure. Start by accepting the standard definition of a relativistic spacetime as a pair $\mathcal{M}, g_{ab}$, where $\mathcal{M}$ is a Hausdorff differentiable manifold and g_{ab} is a Lorentz signature metric defined at every point $p \in \mathcal{M}$.[15] One might contemplate a literal branching of a relativistic spacetime as pictured in Figure 10.1, which shows an upside down "trousers universe" for which the "trunk" bifurcates into two "legs". However, such a contemplation involves a change in the spatial topology and, thus, it runs up against no-go results for topology change. This subsection will review some of these results. The following subsection will discuss various escape options.

RESULT 1 (Tipler, 1977). Let $\mathcal{M}, g_{ab}$ be a compact n-dimensional (Hausdorff) spacetime with boundary consisting of the disjoint union of two compact $(n-1)$-dimensional spacelike manifolds Σ_1 and Σ_2, and let T_{ab} be a stress energy tensor describing the matter fields that are the sources of gravity. Suppose that the spacetime is time orientable. Suppose further that Einstein's gravitational field equations are satisfied by g_{ab}, T_{ab} at every point of $\mathcal{M}$ and that T_{ab} satisfies the weak

[13] "[I]ndeterminism occurs again in the random or chance selection of the branch which becomes actual" (McCall, 1995, p. 156).

[14] 'GTR' is being used here in a sense that includes STR as a special case. In particular, all of the considerations below apply equally to flat spacetimes.

[15] A topological space is said to be *Hausdorff* just in case for any pair of distinct points there are disjoint open neighborhoods. A *differentiable manifold* is a topological space equipped with a differentiable structure that makes it possible to talk of degrees of smoothness or differentiability that go beyond simple continuity. The exact form of differentiability is not relevant here; for sake of definiteness the spacetime manifold may be assumed to be C^∞. The relevant mathematics for relativistic spacetimes is explained in Wald (1994).

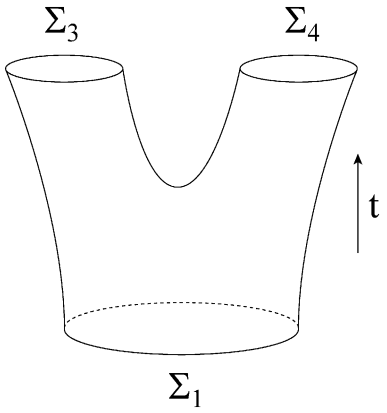

FIGURE 10.1 Trousers spacetime.

energy condition and the generic condition. Then Σ_1 and Σ_2 have the same topology.

A spacetime $\mathcal{M}, g_{ab}$ is said to be *time orientable* iff it admits a non-vanishing, continuous, timelike vector field. Two such fields can be deemed equivalent if at every point $p \in \mathcal{M}$ the vectors they define fall into the same lobe of the null cone at p. The choice of one of two equivalence classes constitutes a *time orientation* for $\mathcal{M}, g_{ab}$. The *weak energy condition* says intuitively that there are no negative energy densities; technically it requires that $T^{ab}V_aV_b \geqslant 0$ for every timelike vector V^a. The *generic condition* says intuitively that every timelike or null geodesic feels a tidal force at some point; technically it requires that every timelike or null geodesic contains at least one point p at which $V^aV^bV_{[c}R_{d]ab[e}V_{f]} = 0$ where V^a is tangent vector at p to the geodesic, R_{abcd} is the Riemann curvature tensor, and the square brackets denote antisymmetrization. Tipler's result can be paraphrased by saying that in a generic time orientable model of GTR with physically reasonable source fields, topology change cannot take place in a spatially finite universe. The power of this result is that it does not appeal to any causality conditions over and above time orientability. For present purposes its weakness lies in the fact that it appeals not only to conditions on the geometry of spacetime but also to physical assumptions, i.e. Einstein's field equations and energy conditions. The other no-go results to be discussed below have the opposite trade off.

In a time oriented spacetime a spacelike hypersurface $\Sigma \subset \mathcal{M}, g_{ab}$ is said to be a *Cauchy surface* iff Σ is intersected exactly once by every future directed timelike curve without past or future endpoint.

RESULT 2 (Geroch, 1967a). Let $\mathcal{M}, g_{ab}$ be a connected time oriented (Hausdorff) spacetime without boundary. Suppose that it admits a Cauchy surface Σ. Then $\mathcal{M}$ is topologically $\Sigma \times \mathbb{R}$.[16]

[16] Further, $\mathcal{M}$ can be foliated by Cauchy surfaces with the topology of Σ; see Wald (1994, Theorem 8.3.13).

Since the trousers universe is not topologically $\Sigma \times \mathbb{R}$ for any Σ, it follows from Result 2 that this universe does not admit a Cauchy surface. Advocates of individual branching who are out to model indeterminism may be unmoved by this result. They can cheerfully give up the existence of Cauchy surfaces, for the absence of such surfaces spells doom for (global) Laplacian determinism.[17] The next result should give them more pause.

RESULT 3 (Geroch, 1967a). Let $\mathcal{M}, g_{ab}$ be a compact n-dimensional (Hausdorff) spacetime with boundary consisting of the disjoint union of two compact $(n-1)$-dimensional spacelike manifolds Σ_1 and Σ_2. Suppose that the spacetime is time oriented and that it contains no closed timelike curves (CTCs).[18] Then Σ_1 and Σ_2 have the same topology.

Apply Result 3 to the trousers universe by taking Σ_1 to be a waist slice, Σ_2 to be the disjoint union of two leg slices Σ_3 and Σ_4 (see Figure 10.1), and $\mathcal{M}$ to be the portion of the trousers universe between and including Σ_1 and Σ_2. Since Σ_1 and Σ_2 have different topologies (in particular, Σ_1 is connected but Σ_2 is not), conclude that this spacetime is either: not time orientable; or is time orientable and contains CTCs; or else $\mathcal{M}$ is non-compact. The last possibility would signal that the spacetime is singular in the sense that the metric is ill-defined somewhere on the portion of the spacetime sandwiched between Σ_1 and Σ_2. Geroch (1967b) and Hawking (1992) generalize Result 3 to the spatially open universes by showing that topology change in bounded regions of a time orientable spacetime entails the existence of CTCs.

A time oriented spacetime $\mathcal{M}, g_{ab}$ is said to be *causally stable* iff (intuitively speaking) it contains no CTCs and, further, there exists a finite widening of the null cones of g_{ab} for which there are no CTCs. The formal definition can be found in Hawking and Ellis (1973, p. 198). But for present purposes the crucial fact is that stable causality is equivalent to the existence of a *global time function*, i.e. a smooth map $t : \mathcal{M} \to \mathbb{R}$ such that $t(p) < t(q)$ for any $p, q \in \mathcal{M}$ such that there is a future directed timelike curve from p to q.

RESULT 4 (Hawking and Ellis, 1973). Let $\mathcal{M}, g_{ab}$ be a stably causal spacetime. If the level surfaces $t = const$ of a global time function are all compact, then they are all topologically the same and $\mathcal{M}$ is topologically $\Sigma \times \mathbb{R}$ for some spacelike hypersurface Σ.

Some additional no-go results are worth mentioning. Gibbons and Hawking (1992a, 1992b) show that existence of a spin structure for spacetime prevents certain kinds of topology change. Alty (1995) shows that for a large class of spacetimes, topology change implies spacetime singularities, whether or not CTCs are

[17] Intuitively, if Σ fails to be a Cauchy surface there will be possible causal processes that fail to register on Σ and, thus, it is unreasonable to expect that even the most precise specification of the state on Σ will suffice to fix, via the laws, the state everywhere in the spacetime.

[18] A CTC is a closed curve whose parametrization can be chosen such that the tangent to the curve is everywhere timelike and future-pointing according to the time orientation.

present. Hawking and Sachs (1974) promote the requirement of *causal continuity*: "There is some reason, but no fully convincing argument, for regarding causal continuity as a basic macrophysical property." Intuitively, this requirement says that the causal past (respectively, future) of any point $p \in \mathcal{M}$—i.e. the portion of spacetime that can be reached from p by a past directed (respectively, future directed causal curve)—should depend continuously on p. Causal continuity is stronger than the requirement of stable causality but weaker than the existence of a Cauchy surface.[19] It is incompatible with the trousers universe of Figure 10.1 and, presumably, with most branching spacetime structures.

3.3 Attempted escapes from the no-go results

Needless to say, all of the no-go results on topology change discussed above rely on substantive assumptions and, thus, can be escaped by rejecting one or more of these assumptions.

Escape 1. Drop the requirement of time orientability. Then none of the no-go results of Section 3.2 is applicable. On the positive side, it is known that dropping time orientability opens the way to topology changes that were not possible with time orientability. As a preliminary stating one precise result, some additional concepts are required.

Two $(n-1)$-dimensional manifolds Σ_1 and Σ_2 are said to be *topologically cobordant* just in case there is an n-dimensional manifold $\mathcal{M}$ with boundary consisting of the disjoint union of the two $(n-1)$-manifolds in question. In the case Σ_1 and Σ_2 are compact it is natural to require that the interpolating $\mathcal{M}$ be compact as well. The necessary and sufficient condition for topological cobordance of compact Σ_1 and Σ_2 is that their Stiefel–Whitney numbers be equal (see Milnor, 1965). It follows that when $n = 4$, any two 3-dim compact Σ_1 and Σ_2 are topologically cobordant. The topologically cobordant manifolds Σ_1 and Σ_2 are said to be *Lorentz cobordant* just in case the interpolating $\mathcal{M}$ admits a time oriented Lorentz signature metric for which Σ_1 and Σ_2 are spacelike. The existence of such a metric is equivalent to the existence of a non-vanishing vector field t^a that is nowhere tangent to Σ_1 or Σ_2. It is natural to require that t^a points into $\mathcal{M}$ on (say) Σ_1 and out of $\mathcal{M}$ on Σ_2. The necessary and sufficient condition for Lorentz cobordance in this sense is that $\chi(\mathcal{M}) = 0$ for even n and $\chi(\Sigma_1) = \chi(\Sigma_1)$ for odd n, where $\chi(\Sigma)$ denotes the Euler characteristic of Σ (see Sorkin, 1986a).

An example of how dropping of time orientability opens up possibilities of topology change is given in Sorkin (1986b). Let Σ_1 and Σ_2 be topologically cobordant $(n-1)$-dimensional compact manifolds. If n is odd then there is a connected manifold $\mathcal{N}$ which has a boundary consisting of the disjoint union of Σ_1 and Σ_2 and which admits a (possibly non-time orientable) Lorentz metric such that Σ_1 and Σ_2 are spacelike.

However, giving up time orientability is not an escape that the advocates of individual branching want to use since they assume that branching is compatible with a consistent time directionality. In addition, Borde (1994) shows that even if

[19] In more detail, the causality conditions discussed above are related as follows: ∃ a Cauchy surface (a.k.a. global hyperbolicity) ⇒ causal continuity ⇒ ∃ a global time function (stable causality) ⇒ no closed CTCs.

time orientability is dropped, topology change implies CTCs[20] if the spacetime is *causally compact*. A (possibly) non-time orientable spacetime $\mathcal{M}$ with boundary consisting of the disjoint union of two spacelike hypersurfaces Σ_1 and Σ_2 is said to be causally compact just in case the topological closure of $I(p)$ is compact for each $p \in \mathcal{M}$, where $I(p)$ consists of all points of $\mathcal{M}$ that can be reached from p by a timelike curve. Intuitively, this condition says that points of $\mathcal{M}$ cannot be causally connected to points at infinity or to "holes" that have been cut in the manifold.

Escape 2. Drop causality requirements, e.g. the requirement that there are no CTCs. Again, this is not an escape that the advocates of individual branching want to use since they assume that branching is compatible with a globally consistent time order, which requires the existence of a global time function.[21] Moreover, this escape does not avoid Result 1 or the result of Alty (1995) neither of which uses causality requirements beyond time orientability.

Escape 3. Go to open universes where the space slices are non-compact. Then Results 2 and 3 do not apply. There are two problems with this escape. First, what if the actual universe is spatially closed? Note that cosmologists use the Friedmann–Robertson–Walker (FRW) models to describe the large scale structure of our universe, and the latest astronomical data is (barely) consistent with a $k = +1$ FRW universe, which has compact space slices. Second, going to spatially open universes does not escape the generalizations of Result 3 proved by Geroch (1967b) and Hawking (1992) if the branching is localized, i.e. is confined to a compact region of space.

Escape 4. Escape by permitting spacetime singularities. This escape can be described in two ways. First, continue to operate with the definition of a relativistic spacetime as a pair $\mathcal{M}, g_{ab}$ where $\mathcal{M}$ is a Hausdorff differentiable manifold, but allow that g_{ab} may not be defined at every point $p \in \mathcal{M}$. Then Results 1 and 3 do not apply even if $\mathcal{M}$ is a compact Hausdorff manifold with boundary, for both results implicitly assume that the metric is everywhere defined on $\mathcal{M}$. On the positive side, it is known that if g_{ab} is allowed to be undefined at (at most) a finite number of interior points of a manifold with boundary, then any two compact three manifolds are Lorentz cobordant; and, moreover, the causal structure is well defined even at the points where the metric is not defined (see Sorkin, 1990 and Borde et al., 1999).[22] Alternatively, one could stick to the original definition of a relativistic spacetime by deleting those interior points of the manifold $\mathcal{M}$ with boundary where the metric is ill-defined, with the result that the originally compact $\mathcal{M}$ is

[20] If the spacetime is not time orientable, then a CTC must be understood as a closed curve whose tangent is everywhere timelike, dropping the condition that the tangent is everywhere future pointing.

[21] However, the advocates of branching spacetimes might want to cite Sorkin's (1986a) interpretation of Result 1 above as showing that causality violations connected with topology changes will be confined to Planck scale regions of spacetime and, therefore are "harmless."

[22] The construction makes use of a *Morse function*. Let $\mathcal{M}$ be a compact n-manifold with boundary consisting of the disjoint union of the $(n-1)$-manifolds Σ_1 and Σ_2. A Morse function is a C^∞ map $f : \mathcal{M} \to [0,1]$ such that $f^{-1}(0) = \Sigma_1$ and $f^{-1}(1) = \Sigma_2$. The critical points of f are those for which $\partial_a f = 0$. These are assumed to be finite in number, and are non-degenerate in the sense that the matrix $\partial_a f \partial_b f$ is non-singular. Since $\mathcal{M}$ is compact it admits a Riemann metric h_{ab}. The associated More metric is then defined by $g_{ab} := (h^{cd}\partial_c f \partial_d f) h_{ab} - \varsigma \partial_a f \partial_b f$, where $\varsigma > 1$ is a real number. This Morse metric vanishes only at the (finite number) of critical points of the Morse function and is Lorentzian everywhere it does not vanish. The version of the construction used in Borde et al. (1999) assigns a causal future and a causal past even to the points where the metric vanishes.

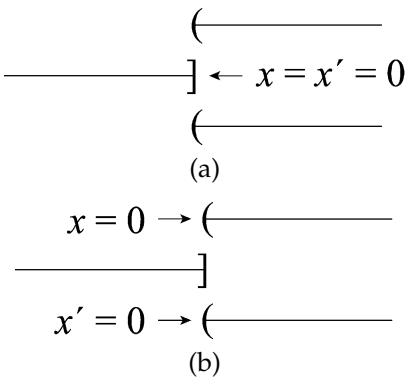

FIGURE 10.2 One-dimensional branching space. (a) The space is non-Euclidean at $x = x' = 0$. (b) The points $x = 0$ and $x' = 0$ are not Hausdorff separated.

no longer compact. Consider the first way of describing the escape. The following dilemma arises. Either there is an isometric embedding ι of $\mathcal{M}, g_{ab}$ into a spacetime $\mathcal{M}', g'_{ab}$ such that any point $p \in \mathcal{M}$ where the metric g_{ab} is not defined has an image point $p' = \iota(p)$ where the metric g'_{ab} is defined, or else not. If the first horn applies, then the no-go results have been evaded by artificially erasing the metric at various points of $\mathcal{M}$ (e.g. the crotch points in the trousers spacetime). Such an artifice ought to be disallowed in constructing models of physically reasonable spacetimes; for physical laws as we know them presuppose a spacetime metric and, thus, the points at which the metric has been erased are literally lawless locations. If the second horn applies, then the points of $\mathcal{M}$ at which the metric is not defined represent genuine singularities in the metric field. The singularity theorems of Hawking and Penrose indicate that spacetime singularities are a generic feature of the solutions of Einstein's field equations;[23] in this sense, classical GTR forces us to seize the second horn in some circumstances. But the advocates of individual branching would be engaging in wishful thinking if they relied on the mechanisms of singularity formation in classical GTR (e.g. gravitational collapse) to create singularities just where they think branching should happen. Furthermore, it is conjectured that quantum gravity effects will smooth out the spacetime singularities of classical GTR.[24]

Escape 5. Allow for spacetimes that are topological spaces but are not locally Euclidean. Such spaces certainly include branching structures, as illustrated for the case of a one-dimensional topological space in Figure 10.2(a) which is constructed as follows. Take two copies of the real line $\mathbb{R}$ coordinatized by x and x', and identify the points of the two copies such that $x = x'$ and $x, x' \leqslant 0$. The resulting tree structure is non-Euclidean at $x = x' = 0$.[25] But such topological spaces cannot be made into spacetimes in the sense of GTR. For topological spaces that

[23] See Wald (1994, Ch. 9) for a review of these results. Most of the theorems use geodesic incompleteness as a criterion of the existence of singularities; geodesic incompleteness may or may not be connected with singularities in a more intuitive sense, e.g. the "blow up" of curvature scalars.

[24] See Earman (2007b) for a review of some of the relevant considerations.

[25] This construction corresponds to McCall's (1994, Appendix 1) "lower cut" option.

are not locally Euclidean cannot be assigned a differentiable structure, and such a structure is essential in formulating the very notion of a Lorentzian metric and in formulating the Einstein field equations.[26] In short, Escape 5 is not viable if one wants to do anything resembling classical GTR.

Escape 6. Keep the standard definition of a relativistic spacetime as a differentiable manifold but abandon the requirement that the manifold be Hausdorff. Now spacetime branching can occur even in spatially closed universes without abandoning temporal orientability, without having to swallow acausalities such as CTCs, and without violating Einstein's field equations, energy conditions, or the generic condition. To get a feel for how dropping the Hausdorff condition allows branching while still allowing the manifold structure needed to formulate GTR, modify the above example by identifying the points of the two copies of $\mathbb{R}$ such that $x = x'$ and $x, x' < 0$ (see Figure 10.2(b)). The result is a locally Euclidean space in which the points $x = 0$ and $x' = 0$ are not Hausdorff separated.[27] Explicit details on the construction of non-Hausdorff spacetimes can be found in Douglas (1995), McCabe (2005), and Visser (1996, 250–255). The non-Hausdorff branching can be global; e.g. take a $k = +1$ FRW model, and for some branch time $t = t_b$ attach non-Hausdorffly a future "leg" to produce a spacetime resembling the trousers universe of Figure 10.1.[28] Or the non-Hausdorff branching might be local; e.g. it could take place along the future light cones of selected spacetime points (see Penrose, 1979 and McCabe, 2005).

Escapes 1–5 have sufficiently high prices that they do seem to merit further consideration, leaving Escape 6 as the only serious contender. But as we will now see, the non-Hausdorff option has many unattractive features.

3.4 Shortcomings of the non-Hausdorff option

The assumption of Hausdorffness is explicitly invoked only sporadically in textbooks on general relativity. But it is implicitly assumed in so many standard results in GTR that dropping it would require a major rewriting of textbooks. Here are two examples of widely used results that depend on Hausdorffness. (i) A compact set of a topological space is closed—if the space is Hausdorff. (ii) If a sequence of points of a topological space converges, the limit point is unique—if the space is Hausdorff. The situation is best summed up by a dictum of Robert Wald in response to my query of what relativistic physics would be like without the assumption that the spacetime manifold is Hausdorff: "Asking what relativistic

[26] The metric g_{ab} is a bilinear map from pairs of tangent vectors of the manifold $\mathcal{M}$ to the real numbers. Einstein's field equations are formulated in terms of second-order derivatives of the metric, the definition of which relies on the differentiable structure of $\mathcal{M}$.

[27] This construction corresponds to McCall's (1994, Appendix 1) "upper cut" option.

[28] Here t is the global time function for the FRW spacetime metric as displayed in the line element

$$ds^2 = a(t)\left[\frac{dr^2}{1-kr^2} + r^2\,d\theta^2 + r^2\sin^2\theta\,d\varphi^2\right] - dt^2,$$

where $a(t)$ is the scale factor (sometimes called the radius of the universe) and $k = 0, -1$, or $+1$, corresponding respectively to space sections of zero curvature, constant negative curvature, and constant positive curvature.

physics would be like without Hausdorffness is like asking what the earth would be like without its atmosphere."[29]

Specific illustrations of difficulties for relativistic physics that would crop up in non-Hausdorff spacetimes are easy to produce. For instance, non-Hausdorff spacetimes can admit bifurcating geodesics; that is, there can be smooth mappings γ_1 and γ_2 from, say, $[0,1] \subset \mathbb{R}$ into $\mathcal{M}$ such that the image curves $\gamma_1[0,1]$ and $\gamma_1[0,1]$ are geodesics that agree for $[0,b)$, $0 < b < 1$, but have different endpoints $\gamma_1(1)$ and $\gamma_2(1)$.[30] This does not contradict a fundamental theorem for both Riemannian and Lorentzian manifolds which asserts the *local* existence and uniqueness of geodesics: consider any point $p \in \mathcal{M}$ and any (non-zero) vector $V_p \in T_p(\mathcal{M})$; then for some interval $-\delta < t < +\delta$, $\delta > 0$, there is a unique geodesic $\gamma(t)$ such that $\gamma(0) = p$ and $(d\gamma/dt)_{t=0} = V_p$. But in the presence of non-Hausdorff branching this local theorem cannot be used to conclude that each geodesic is contained in a unique maximal geodesic. According to GTR, the worldline of a massive test particle not acted upon by non-gravitational forces is a timelike geodesic. But how would such a particle know which branch of a bifurcating geodesic to follow? This problem has led general relativists to shun non-Hausdorff spacetimes that involve non-Hausdorff branching (see Hajicek 1970, 1971; and Miller, 1973).[31]

Another example concerns the failure of either local or global conservation laws. Local differential geometry can be developed per usual on non-Hausdorff spacetimes. In particular, one can impose the local conservation law $\nabla_a T^{ab} = 0$ for the stress energy-tensor T^{ab}.[32] The implicit assumption is that T^{ab} is continuous and, indeed, differentiable. But this assumption can be violated when a branch is non-Hausdorffly glued on unless T^{ab} is smoothly extended along the branch. Thus, to the above rhetorical question, "Which branch of a bifurcating geodesic should a free-falling test particle follow?", the answer should be "All!" if local conservation is to hold.

So suppose then that T^{ab} is smoothly extended along all branches. Then global conservation laws will be violated. To see this, consider how the local conservation law $\nabla_a T^{ab} = 0$ for the stress energy-tensor T^{ab} is normally integrated to give a global conservation law. Suppose that the spacetime $\mathcal{M}, g_{ab}$ is stationary, i.e. admits a non-vanishing timelike vector field V^a such that $\nabla_{(a} V_{b)} = 0$.[33] Define the momentum flow (associated with V^a) as $P^a := -T^{ab} V_b$; then $\nabla_a P^a = 0$. Suppose finally that T^{ab} satisfies the dominant energy condition;[34] then P^a is a future directed non-spacelike vector. Then using the spacetime version of Gauss' theorem it can be shown that total energy is conserved:

[29] Private communication to the author.

[30] Non-Hausdorffness is a necessary but not sufficient condition for the existence of bifurcating geodesics; see Hajicek (1971) where a necessary and sufficient condition for a non-Hausdorff spacetime to have no bifurcating geodesics is given. But such bifurcation is a product of the structure advocates of individual branching want and need to escape the no-go results on topology change given above.

[31] Of course, those who want to use branching in individual spacetimes to express indeterminism may like bifurcating geodesics. They should read on.

[32] ∇^a is the covariant derivative determined by the metric g_{ab}.

[33] This is the invariant way of saying that the metric is not time dependent.

[34] This condition, which is needed to prove that there is no superluminal transmission of energy-momentum, requires that for any timelike vector V^b, $-T^a_b V^b$ is a future directed timelike or null vector, is thought to be satisfied by all of the matter fields encountered in classical GTR; see Wald (1994, p. 219).

LEMMA. Let $\mathcal{M}, g_{ab}$ be a compact n-dimensional (Hausdorff) spacetime with boundary consisting of the disjoint union of two compact $(n-1)$-dimensional spacelike manifolds Σ_1 and Σ_2. Under the above assumptions about P^a, the energy $E(\Sigma_1) := \int_{\Sigma_1} P^a \, dS_a$ at "time" Σ_1 equals the energy $E(\Sigma_2) := \int_{\Sigma_2} P^a \, dS_a$ at "time" Σ_2.

This lemma on the conservation of energy can be generalized to the case of non-compact space slices when the support of P^a is confined to a finite world tube. But this lemma and its generalizations can fail for non-Hausdorff branching spacetimes if T^{ab} is smoothly extended along the branch—the total energy can increase with time. If "observers" are associated with non-branching segments of (possibly) branching worldlines, then no observer will ever be directly aware of a violation of conservation of energy. But an observer who believes in individual branching will have reason to believe that it occurs.

The non-Hausdorff modeling of spacetime branching also suffers from the inability to accommodate various ideas. (i) The idea that the laws of physics plus what happens up to and including a given moment of time do not uniquely fix the future. The non-Hausdorff branching can (at best) capture the idea that what happens up to but not including a given moment of time does not uniquely fix the future. To include the given moment would seem to require non-locally Euclidean spacetimes. (ii) Indeterminism could involve a radical "openness" of the future in the sense that, compatible with the laws of physics and the past history, there are possible global alternative at *each and every* future moment of time. But, consistent with spacetime being a manifold, this cannot be modeled in terms of non-Hausdorff branching at every moment. Note that ensemble branching has no trouble in expressing the ideas in (i) and (ii).

As a final drawback of non-Hausdorffness I will mention the lack of control that results from allowing non-Hausdorff manifolds. As noted in Section 3.1, branching spacetimes do not in themselves entail indeterminism in the proper sense of non-uniqueness of dynamical evolution. But allowing non-Hausdorff branching does undermine the determinism of classical GTR. If Hausdorffness is assumed, then GTR provides for a (locally in time) deterministic evolution: the basic theorem on the initial value formulation of Einstein's field equations shows that corresponding to appropriate initial data on a three manifold (thought of as specifying the state at a given instant), there exits a unique (up to diffeomorphism) maximal solution for which the initial value hypersurface is a Cauchy surface (see Theorem 10.2.2 of Wald (1994)).[35] The uniqueness result fails if non-Hausdorff branching is allowed. The failure may be applauded by advocates of individual branching. But general relativists will be appalled by the lack of control on the addition of non-Hausdorff branches. Suppose for sake of illustration that the large scale structure of our universe is captured by a FRW big bang model. Start with the standard Hausdorff version of this model, and at some time t_b attach non-Hausdorffly n branches. What determines the choice of t_b and n? Nothing in the laws of classical GTR or any natural extension thereof provides an answer.

[35] This theorem says nothing about how "big" the unique solution is or whether it is maximal simpliciter. These issues are connected with the problem of cosmic censorship.

The point here is not peculiar to non-Hausdorff branching spacetimes but applies equally to other constructions that are explicitly or implicitly shunned by physicists. For example, suppose that the triple $\mathcal{M}, g_{ab}, T_{ab}$ satisfies Einstein's gravitational field equations and that T_{ab} satisfies standard energy conditions (say, the weak and dominant energy conditions). Choose any closed set of points $\mathcal{C} \subset \mathcal{M}$. Then surgically removing this set of points results in a triple $\mathcal{M} - \mathcal{C}, g_{ab}|_{\mathcal{M}-\mathcal{C}}, T^{ab}|_{\mathcal{M}-\mathcal{C}}$ that also satisfies Einstein's field equations and energy conditions. Conspiracy theorists who want to explain strange disappearances will revel in this construction. Don't bother to look in Florida or California for the children whose pictures appear on milk cartons. In fact, don't bother to look for them anywhere in spacetime; for their world lines have fallen into "holes" carved into the spacetime (i.e. take $\mathcal{C}$ to be the disjoint union of closed balls of $\mathcal{M}$). Creationists will also be delighted. To model how the universe could have begun at 10,000 B.C., take $\mathcal{C}$ to consist of all of those points on or to the past of a spacelike slice that corresponds to 10,000 years ago. The result is a universe that pops into existence with the fossils and all of the other physical artifacts that the deluded proponents of evolution take as evidence for their theory. Since classical GTR has no way to control the pathologies such models entail, they must be excluded in a sensible version of the theory. And, indeed, it is explicitly or implicitly assumed by general relativists that physically reasonable models of spacetime should be maximal, i.e., cannot be isometrically embedded as a proper subset of a larger spacetime.[36] Non-Hausdorff branching spacetimes are, I submit, in exactly the same league with "holey" spacetimes.

If non-Hausdorff spacetime branching is to be taken seriously, what is needed is a physical theory that prescribes the dynamics of branching. What the advocates of individual branching offer is not a theory but the nostrum that branching takes place when an indeterministic process produces an outcome plus *ad hoc* rules, e.g. there is branching to the future but not towards the past. This is no more satisfactory than saying in QM that a reduction of the state vector takes places when a measurement occurs. Unlike other theoretical terms employed by the theory, "measurement" is not a term that can play a role in explanations; it is rather something that requires analysis and explanation. Similarly, "outcome of an indeterministic process" is not a term of any theory of physics but something that needs analysis. Note that I am not criticizing state vector reduction *per se*. There are honest theories of state vector reduction that offer a dynamics that explains how and when state vector reduction takes place (see, for example, Pearle 1976, 1989). I know of no extant honest theories of how and when non-Hausdorff branching takes place.

Where might one look for a dynamics of non-Hausdorff branching? Classical GTR is a theory that provides a dynamics for spacetime; but it eschews intrinsic randomness. The situation is just the opposite for QM and quantum field theory: these theories deal with intrinsic randomness, but they operate on a fixed spacetime background. The Holy Grail of current physics research is a quantum theory of gravity—a theory that combines the insights of GTR and quantum physics.

[36] Whether or not this condition is sufficient to exclude "holey" spacetimes is an interesting issue, but it is not one that needs to be discussed here.

In one sense, quantum gravity will probably render moot the issue of branching spacetime since at a fundamental level the spacetime of classical GTR will dissolve into quantum foam. But one can ask: In those circumstances when a classical general relativistic spacetime emerges in some semi-classical limit of quantum gravity, will the emergent spacetime be a non-Hausdorff branching spacetime? My conjecture is that the answer is negative. Some support for this conjecture comes from the finding that quantum fields display singular behavior when propagating on a $(1+1)$-dimensional trousers spacetime (see Anderson and DeWitt, 1986; Manogue et al., 1988). This has led to the opinion that the kind of topology change involved in a trousers spacetime, and more generally in any spacetime that violates causal continuity, will be suppressed in the sum over histories form of quantum gravity (see Anderson and DeWitt, 1986; Dowker, 2003).[37]

In summing up the discussion of Escape 6, I agree with Roger Penrose, who after a brief consideration of non-Hausdorff branching spacetimes, concluded: "I must ... return firmly to sanity by repeating to myself three times: 'spacetime is a *Hausdorff* differentiable manifold; spacetime is a *Hausdorff* ...'!" (1979, p. 595).

4. CONCLUSION

Some of the ways of implementing branching in individual spacetimes have what strike me as fatal defects. None of the objections to non-Hausdorff branching can be deemed to be fatal, but the cumulative weight of these objections seems to me justify the attitude that non-Hausdorff branching is to be viewed as a desperate last resort rather than a device to be used as the basis for a program aimed at interpreting scientific or ordinary language concepts. Overall, the recommendation to the proponents of the many worlds interpretation of QM or the Belnap school of branching spacetimes is to proceed without recourse to individual branching.

Regardless of the merits of my stance, the study of branching spacetimes reveals itself as a fruitful way to bring together issues from philosophy (e.g. understanding the "openness" of the future), general philosophy of science (e.g. the analysis of determinism), and the foundations of physics (e.g. the role of branching spacetimes in classical GTR and quantum gravity).

ACKNOWLEDGEMENT

I am grateful to Gordon Belot, Craig Callender, Robert Geroch, John Norton, and Steve Savitt for helpful suggestions on a earlier draft of this chapter.

REFERENCES

Alty, L.J., 1995. Building blocks for topology change. Journal of Mathematical Physics 36, 3613–3618.

[37] The goal of the sum over histories approach is to calculate transition amplitudes between two n-geometries representing the states of an $(n+1)$-dimensional spacetime at two different times. This is to be accomplished by summing over a weighted average of the actions associated with all of the possible histories connecting the two states.

Anderson, A., DeWitt, B., 1986. Does the topology of space fluctuate? Foundations of Physics 16, 91–105.
Belnap, N., 1992. Branching space-time. Synthese 92, 385–434.
Belnap, N., 2002. EPR-like 'funny business' in the theory of branching space-times. In: Placek, T., Butterfield, J. (Eds.), Non-Locality and Modality. Kluwer Academic, Dordrecht, pp. 293–315.
Belnap, N., 2003. No-common-cause EPR-like 'funny business' in the theory of branching space-times. Synthese 114, 199–221.
Belnap, N., Perloff, M., Xu, M., 2001. Facing the Future: Agents and Choices in Our Indeterministic World. Oxford University Press, Oxford.
Borde, A., 1994. Topology change in classical general relativity. gr-qc/9406053.
Borde, A., Dowker, H.F., Garcia, R.S., Sorkin, R.D., Surya, S., 1999. Causal continuity in degenerate spacetimes. Classical and Quantum Gravity 16, 3457–3481.
Broad, C.D., 1923. Scientific Thought. Harcourt, Brace, & Company, New York.
Douglas, R., 1995. Stochastically branching spacetime topology. In: Savitt, S. (Ed.), Time's Arrow Today. Cambridge University Press, Cambridge, pp. 173–188.
Dowker, F., 2003. Topology change in quantum gravity. In: Gibbons, G.W., Shellard, E.P.S., Rankin, S.J. (Eds.), The Future of Theoretical Physics and Cosmology: Celebrating Stephen Hawking's 60th Birthday. Cambridge University Press, Cambridge, pp. 436–452.
Earman, J., 2006. The implications of general covariance for the ontology and ideology of spacetime. In: Dieks, D. (Ed.), The Ontology of Spacetime I. Elsevier, Amsterdam.
Earman, J., 2007a. Aspects of determinism in modern physics. In: Butterfield, J., Earman, J. (Eds.), The Philosophy of Physics. Handbook of the Philosophy of Science, vol. 2. Elsevier/North-Holland, Amsterdam.
Earman, J., 2007b. Essential self-adjointness: Implications for determinism and the classical-quantum correspondence. Synthese, to appear.
Geroch, R.P., 1967a. Topology in general relativity. Journal of Mathematical Physics 8, 782–786.
Geroch, R.P., 1967b. Singularities in the space-time of general relativity: Their definition, existence, and local characterization. PhD dissertation, Princeton University.
Gibbons, G.W., Hawking, S.W., 1992a. Kinks and topology change. Physical Review Letters 69, 1719–1721.
Gibbons, G.W., Hawking, S.W., 1992b. Selection rules for topology change. Communications in Mathematical Physics 148, 345–352.
Hajicek, P., 1970. Extension of the Taub and NUT spaces and extensions of their tangent bundles. Communications in Mathematical Physics 17, 109–126.
Hajicek, P., 1971. Bifurcate space-time. Journal of Mathematical Physics 12, 157–160.
Hawking, S.W., 1992. Chronology protection conjecture. Physical Review D 46, 603–611.
Hawking, S.W., Ellis, G.F.R., 1973. The Large Scale Structure of Space-Time. Cambridge University Press, Cambridge.
Hawking, S.W., Sachs, R.K., 1974. Causally continuous spacetimes. Communications in Mathematical Physics 35, 287–296.
Kowalski, T., Placek, T., 1999. Outcomes in branching space-time and GHZ–Bell theorems. British Journal for the Philosophy of Science 50, 349–375.
Lewis, D.K., 1986. On the Plurality of Worlds. Blackwell, Oxford.
MacFarlane, J., 2003. Future contingents. Philosophical Quarterly 53, 321–336.
Manogue, C.A., Copeland, E., Dray, T., 1988. The trousers problem revisited. Pramana 30, 279–292.
McCabe, G., 2005. The topology of branching universes. Foundations of Physics Letters 18, 665–676.
McCall, S., 1994. A Model of the Universe. Clarendon Press, Oxford.
McCall, S., 1995. Time flow, non-locality, and measurement in quantum mechanics. In: Savitt, S. (Ed.), Time's Arrow Today. Cambridge University Press, Cambridge, pp. 155–172.
McCall, S., 2000. QM and STR: The combining of quantum mechanics and relativity theory. Philosophy of Science 67, S535–S548.
Miller, J.G., 1973. Global analysis of the Kerr–Taub–NUT metric. Journal of Mathematical Physics 14, 486–494.
Milnor, J.W., 1965. Topology from a Differentiable Viewpoint. University of Virginia Press, Charlottsville, VA.

Müller, T., 2004. Probability theory and causation: A branching spacetime analysis. http://philsci-archive.pitt.edu.
Pearle, P., 1976. Reduction of the state vector by a nonlinear Schrödinger equation. Physical Review D 13, 857–868.
Pearle, P., 1989. Combining stochastic dynamical state-vector reduction with spontaneous localization. Physical Review A 39, 2277–2289.
Penrose, R., 1979. Singularities and time-asymmetry. In: Hawking, S.W., Israel, W. (Eds.), General Relativity: An Einstein Centenary Survey. Cambridge University Press, Cambridge, pp. 581–638.
Placek, T., 2000. Stochastic outcomes in branching space-time: Analysis of Bell's theorems. British Journal for the Philosophy of Science 51, 445–475.
Placek, T., Müller, T., 2005. Counterfactuals and historical possibility. http://philsci-archive.pitt.edu.
Sorkin, R., 1986a. Topology change and monopole creation. Physical Review D 33, 978–982.
Sorkin, R., 1986b. Non time-orientable Lorentz cobordism calls for pair creation. International Journal of Theoretical Physics 25, 877–881.
Sorkin, R.D., 1990. Consequences of spacetime topology. In: Cooley, A., Cooperstock, F., Tupper, B. (Eds.), Proceedings of the 3rd Canadian Conference on General Relativity and Relativistic Astrophysics. World Scientific, Singapore, pp. 137–163.
Tipler, F.J., 1977. Singularities and causality violation. Annals of Physics 108, 1–36.
Visser, M., 1996. Lorentzian Wormholes: From Einstein to Hawking. AIP Press, Woodbury, NY.
Wald, R.M., 1994. General Relativity. University of Chicago Press, Chicago.

CHAPTER 11

Time Lapse and the Degeneracy of Time: Gödel, Proper Time and Becoming in Relativity Theory

Richard T.W. Arthur*

Abstract

In Special Relativity time bifurcates into coordinate time and proper time. I argue that confusion between these two concepts is largely responsible for the (fallacious) argument that becoming is incompatible with relativity theory.

1. PREVIEW OF THE ARGUMENT

In the transition to Einstein's theory of Special Relativity (SR), certain concepts that had previously been thought to be univocal or absolute properties of systems turn out not to be. For instance, mass bifurcates into (i) the relativistically invariant *proper mass* m_0, and (ii) the mass relative to an inertial frame in which it is moving at a speed $v = \beta c$, its *relative mass* m, whose quantity is a factor $\gamma = (1 - \beta^2)^{-1/2}$ times the proper mass, $m = \gamma m_0$.

By an extension of a term already used in physics, I call this phenomenon *degeneracy*: Just as an energy state is considered degenerate if it is in fact a multiplicity of energy states that are not distinguishable from one another until, say, a magnetic field is applied, so too it turns out that the concept of mass is degenerate: it is only at speeds that are an appreciable fraction of the speed of light that the two different concepts of mass are distinguishable. In the same way one can say that Special Relativity shows that the concept of time is likewise degenerate:

(i) there is the *relative time (or time co-ordinate function)* t, whose quantity varies (like relative mass) according to the inertial frame chosen, and
(ii) there is the *proper time* τ which is invariant under change of frame, and calculated by an integration along the path taken.

* Department of Philosophy, McMaster University, Hamilton, Canada

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00011-9
© Elsevier BV
All rights reserved

This much is well known. But what I think has been insufficiently appreciated is the change in the ontology of time accompanying this degeneracy, especially with respect to issues concerning becoming and process. For the classical concept of time does double duty in

(a) correlating distant events as prior to, simultaneous with, or after, some given event; and
(b) measuring or determining how fast things age, that is how fast the properties of a given system change, or how fast the states of a given process follow one another.

But whereas classically it was assumed that events occur (or states successively come into existence) in the present, where each successive present is assumed to be a unique worldwide instant, this conception is untenable in relativistic physics because of the relativity of simultaneity. That is, if one accepts that the becoming of an event at a particular point in spacetime should be invariant, yet construes its becoming in terms of coming to be simultaneous with a given observer, then one can derive a contradiction. This circumstance has persuaded many physicists and philosophers, pre-eminent among them Kurt Gödel, that one must abandon the idea of becoming altogether. I shall argue that such arguments are fallacious precisely because they keep fused together the two distinct functions of time that are (degenerately) fused in the classical conception of time: (a) determining a plane of simultaneity, and (b) invariant time lapse. In SR the becoming of events in succession, the rate of a process or the rate at which a thing ages, is tracked by *proper time*; the synchronicity of distant events is tracked by the *time-coordinate function*. This separation of these two different aspects of time into two different time concepts τ and t is characteristic not only of SR, but of all relativistic physics, where every timelike curve represents a possible process, whose rate of evolution is parametrized by proper time.

This connection of becoming or process with proper time as opposed to coordinate time is masked, so I argue, by the way many physicists and philosophers talk about proper time as if it is simply the *relative time of an observer in his or her own rest frame*. This is particularly evident in some discussions of the Twin Paradox of SR. By paying careful attention to what the twins could actually observe and infer about each other's times, I argue that it is implicit in the correct resolution of the paradox that time lapse is correctly measured by proper time, and show how, *contra* Gödel, this involves a path-dependent and local notion of becoming. I show how, despite the bifurcation of length in relativistic contexts into frame-dependent and proper length concepts, this has no parallel ontological significance, and briefly discuss this disanalogy between proper time and proper length. In closing, I argue that the detachment of objective becoming from the issue of the existence of invariant time slices undercuts Gödel's arguments for the ideality of time, and discuss the implications of the fact that a timelike curve must represent the path of a possible process.[1]

[1] A complementary analysis of Gödel's arguments, from very much the same point of view as that offered here, is given by Dennis Dieks in his (2006).

2. GÖDEL'S ARGUMENTS FOR THE UNREALITY OF TIME

In 1949 Gödel published his famous solution to the Einstein Field Equations for the General Theory of Relativity, representing a rotating universe in which there exist closed time-like curves (1949a). That is, in such worlds for any two points P and Q on a worldline of matter occurring in the solution, with P preceding Q on this line, "there exists a [continuously future-directed] timelike line connecting P and Q on which Q precedes P; i.e. it is theoretically possible in these worlds to travel into the past, or otherwise influence the past" (447). As is well known, in the same year he published a second paper drawing out the philosophical implications of this, and promoting an idealist philosophy of time (1949b). Gödel argued that for every assignment one might make of a cosmic time function for such models of the universe, "one could travel into regions of the universe which are past according to that definition," thus showing that "an objective time lapse would lose every justification in these worlds" (561). This being so, the assumption of an objective time lapse should be abandoned altogether: time, as Parmenides and Kant had argued, is bound up with our particular way of perceiving the world, and is not an objective feature of the four-dimensional totality.

Less remarked upon is the argument Gödel gives for the same idealistic conclusion earlier in that paper, based on the Special Theory of Relativity. He argues that the assertion that two spatially separated events A and B are simultaneous "loses its objective meaning, insofar as another observer, with the same claim to correctness, can assert that A and B are not simultaneous (or that B happened before A)" (557). This, he claims, allows one to construct "an unequivocal proof for the view of those idealistic philosophers who, like Parmenides, Kant and the modern idealists, deny the objectivity of change and consider change as an illusion or an appearance due to our special mode of perception" (557). The proof he gives runs as follows:

> Change becomes possible only through the lapse of time. The existence of an objective lapse of time, however, means (or at least is equivalent to the fact) that reality consists in an infinity of layers of "now" which come into existence successively. But, if simultaneity is something relative in the sense just explained, reality cannot be split up into such layers in an objectively determined way. Each observer has his own set of "nows", and none of these various systems of layers can claim the prerogative of representing the objective lapse of time. (557–558)

Gödel's idealistic conclusion is of course a radical one, and few modern philosophers of science accept it. According to the dominant view in the philosophy of science in the latter half of the last century, time intervals as measured in the various possible frames of reference are all perfectly objective, even if they are not invariant. Relativity of the duration of a process, it is argued, no more entails its subjectivity or illusory nature than relativity of the mass of a system to frame of reference entails the subjectivity or illusory nature of mass. The relativity of simultaneity entails that no one of these relative times can be privileged as the "actual time", just as Gödel had argued. Nevertheless, each measure of duration

is consistently related to any of the others by the Lorentz transformation formulas. According to the dominant view—as subscribed to, for instance, by Jack Smart, Adolf Grünbaum, Paul Davies, and many others[2]—what is refuted by such arguments from the relativity of simultaneity is not the objectivity of time lapse, but the notion of *coming into existence*. It is true, as Gödel observed, that different choices of inertial reference frame will result in wholly different classes of events being simultaneous with a given event, and that one must therefore relinquish the classical notion of a world-wide "now". But what this precludes is not the objectivity of time lapse, but—as Hilary Putnam (1967) and C.W. Rietdijk (1966) each argued on grounds similar to Gödel's in the late 60's—any notion of *objective becoming* (*becoming real*, in Putnam's case, *becoming determined* in Rietdijk's). (A similar argument was given by Nicholas Maxwell (1985).) Thus although time lapse is perfectly objective, it is frame-dependent.

Interestingly, Gödel, anticipated this objection that the relativity of time lapse "does not exclude that it is something objective." To this he countered that the lapse of time connotes "a change in the existing", and "the concept of existence cannot be relativized without destroying its meaning completely" (558, n. 5). The dominant view, by contrast, would urge that the relativity of existence is avoided precisely by denying that time lapse constitutes a "change in the existing": the existence of events is their existence in a four-dimensional spacetime, and this does not change. Against this, I have argued elsewhere (Arthur, 2006, 131–136) that the sense in which spacetime "exists" is not a temporal sense, and so will not support the contention of Putnam et al. that events simultaneous with another event are "already real" for it; to suppose that this is so, I argue, leads inexorably to a conclusion that denies the reality of temporal succession.[3]

What I wish to draw attention to here, however, is a premise that the dominant view shares with Gödel's: both assume that events are real or determined when they are present to an observer, with presentness construed in terms of simultaneity in the observer's frame of reference; i.e. they construe the reality of an event *in terms of the time co-ordinate function*. Thus Putnam (1967) and Rietdijk (1966) each assume that becoming real or determined must occur relative to "an observer's inertial system", with time-lapse measured by the time co-ordinate function, as a premise in their *reductio* arguments against the reality of becoming real or determinate. The crucial premise here is the Gödelian one that for each individual observer, "the existence of an objective lapse of time … is equivalent to the fact that reality consists in an infinity of layers of 'now' which come into existence successively." That is, the time lapse between, for instance, two events in anyone's life history is given by the difference in the values of the time co-ordinate function in some particular inertial reference frame.

But this construal of time lapse in SR is false, as can be shown by an analysis of the much discussed Twin Paradox. Here we imagine one twin staying at home

[2] See Smart (1968, 255ff.) and (1980); Grünbaum (1976); Davies (1989, 3): "Thus relativity physics has shifted the moving present out from the superstructure of the universe, into the minds of human beings, where it belongs … present day physics makes no provision whatever for a flowing time …".

[3] For a thorough analysis of the problematic notion of *existence* in a temporal context see Steve Savitt (2006) and Mauro Dorato's article (2006) in the same volume.

while the other speeds off at a relative velocity which is an appreciable fraction of *c*, the speed of light, turns round, and returns at a similar velocity. When they are reunited, less time has elapsed for the travelling twin, who is consequently found to have aged less. But the discrepancy between the times elapsed for the two twins cannot be a discrepancy between times as measured by *co-ordinate time*—the time or "layer of 'now'" associated with some given inertial system—since in that inertial frame of reference the twins are apart for exactly the same time, as measured by the time co-ordinate of that frame. Indeed, in any such inertial frame, there is only one difference between the co-ordinates of these two points, and not one for each twin. In fact, the time taken for the twins to make each of their trips through spacetime from the point at which the travelling twin departed to the later point of their reunion must instead be determined by integrating the *proper time* along each twin's particular world line. Thus the root of the trouble with the "layer of now" conception of time lapse is a failure to take into account the degeneracy of time. Time lapse is measured by the proper time. The difference in the proper times for their journeys is not the same as the difference in the time co-ordinates of the two points in some inertial reference frame. If time lapse were measured by such a time co-ordinate function, then both twins would be the same age. They are not. *Ergo*, time lapse (in the sense of how long a given process takes, how quickly it becomes) is not measured by the time co-ordinate function. So Gödel's "unequivocal proof" of the ideality of time falls flat on its face.[4]

It is puzzling that this simple consideration is not widely recognized; this suggests that there are other assumptions at work that mask its application. I believe they have to do with a *misconception of proper time as the time co-ordinate of the observer's rest frame*, and related misconceptions about *an observer "inhabiting an inertial frame"* and "*experiencing*" the events which are simultaneous with his or her state of consciousness. Rietdijk, for example talks of two spatially separated observers "experiencing the same present … in virtually the same inertial system" (1966, 342), Grünbaum writes that for an organism *M* to experience an event at a time *t* is to be "*conceptually aware* of experiencing at that time either the event itself or another event simultaneous with it in *M*'s reference frame" (1976, 479), and Putnam of "everything that is simultaneous to you-now in *your* co-ordinate system" being real, and Clifton and Hogarth of two observers' "inhabit[ing] the same inertial frame" (1995, 379). Although misconceptions about proper time are seldom stated explicitly, they also appear to be quite prevalent. Indeed, they afflict the understanding of SR itself, as witnessed by some of the attempted resolutions of the Twin Paradox.

These considerations motivate another look at the Twin Paradox, to get clear on what is in an observer's (visual) experience in a relativistic context, and what is inferred; and to see more clearly how the distinction between proper time and co-ordinate time cleanly resolves the paradox without reference to the events one

[4] It may be countered, as by my anonymous referee, that Gödel's argument depends only on the lapses of time being different for any two arbitrary curves connecting two timelike related events, and that Gödel does not assume that time lapse is measured by a time-coordinate function. But Gödel explicitly construes time lapse in terms of co-ordinate time in his argument from Special Relativity, where his argument against the "relativization of existence" crucially depends on this. This is supported by the interpretation of Palle Yourgrau (1991), who construes Gödel's argument as depending on a conception of time lapse as relative to reference frame.

"experiences" as present undergoing a dramatic change, (Section 3) or implying that the discrepancy in the twin's ages is a General Relativistic effect (Section 4). This may seem otiose, given the number of times the paradox has been resolved, and given that no one who knows relativity thinks it a problem. But anyone who believes that the resolution of the paradox requires General Relativity, or a recognition that the events experienced as present by the moving twin undergo a discontinuous shift at the point of return, or that there is any asymmetry in what speeding up or slowing down of clocks is seen by or inferred about either twin of the other, has not properly appreciated, so I would claim, the profundity of the changes in our understanding of time wrought by Special Relativity.

3. THE TWIN PARADOX REVISITED

To make our well-worn paradox vivid, let us take Terence the true Tellurian, who tethers himself to terra firma; and Astrid the astronaut, who abridges her age by abandoning Earth with alacrity for Alpha Centauri. (Since the tale of the two twins has been told so many times, I hope I may be allowed a little alliteration in the account.) We'll assume, to keep the figures round, that Alpha Centauri is 6 light years away, and that Astrid approaches it at six tenths of the speed of light ($0.6c$), turns round in an instant, and returns towards Earth at the same speed (all of this in Terence's rest frame, i.e. from the point of view of an inertial frame of reference in which Terence is stationary). An easy calculation shows that, according to Terence, his twin is away for exactly twenty years (ignoring for now any periods of acceleration or deceleration). Things are otherwise for Astrid. At such great speed, distances are foreshortened by a factor $\sqrt{(1-0.36)} = \sqrt{0.64} = 0.8$: it appears to her that she makes a journey outwards of only 4.8 light years, and does it in 8 years. (Seen from Terence's perspective, time runs more slowly on her watch.) An equal space contraction and time dilation occurs on her way home, so that she takes only 16 years for her journey (still discounting the time of her deceleration). Thus when they are reunited and compare their watches, they find that Terence has aged 4 years more than his sister! This seems obviously paradoxical. If all inertial motion is relative, how can there be an absolute difference in their lifetimes resulting from it? The paradox is heightened by this observation: while the twins are in relative inertial motion, each's duration will be running slower from the perspective of the other's rest frame. In each leg of the journey, Astrid would infer processes to be happening more slowly on Earth as it receded from her or approached her at $0.6c$: her eight years would correspond to an inferred duration of processes on earth of only 6.4 years! As Herbert Dingle reasoned,[5] if each twin's life-process is slowed down relative to the other, each will age less than the other, an obvious impossibility! This is the paradox.

A standard resolution given of this paradox explains that the reason for the discrepancy is that as it is only Astrid who undergoes an acceleration as she turns around, it is therefore she who performs an absolute motion, not Terence. On this

[5] Herbert Dingle (1890–1978) was an English astronomer who wrote a standard textbook on relativity theory before becoming a vociferous opponent in his old age. See his (1972).

analysis, so long as the twins are in inertial motion relative to one another, each twin must indeed infer that the other's clock is running slow. The reason for the discrepancy is that, as Astrid turns about, the act of her deceleration skews her temporal orientation violently, and under the conditions stated, instantaneously. As she journeys home she infers Terence's clock to advance only 6.4 years, yet it will read 20 when she returns to Earth. So, at the time she sets out from Alpha Centauri for the journey home, it must read 13.6. Yet the instant before, the instant she arrives at Alpha Centauri, she would have inferred it to have read 6.4 years! So, instantaneously, it would have had to have jumped 7.2 ($= 13.6 - 6.4$) years.

So it's not that Astrid's instantaneous (and wholly unphysical) acceleration introduces a time dilation; it's that it discontinuously skews her temporal orientation. Now if we were to follow Putnam's informal way of speaking, we would say that Astrid "experiences" 7.2 years going by in an instant: events that were "present according to her co-ordinate system" are discontinuously displaced 7.2 years into the past of "her-now" according to that same system. In actual fact, however, no such wrenching change of her experience of the present occurs. These are facets of a sloppy use of the ideas of "observer's reference frame" and the observer's "present", and a failure to distinguish between the time an observer might *infer* an event to occur from when the observer would *see* it occurring.[6] In fact, it will be worth going over the whole thing in some detail to see how this could be the case. Taking my cue from the lucid account of the twin paradox given by Paul Davies in his recent book, I shall re-examine the thought experiment by conducting it in three stages.[7]

First, just to get our bearings, let us assume a Cartesian cosmos, in which light is a *pression* that is transmitted instantaneously, and durations are completely independent of the state of motion of the enduring thing. We equip each twin with a very powerful telescope and a very large clock, and then suppose Astrid to leave for Alpha Centauri at $0.6c$. This is very straightforward. The Tellurian twin sees his astronaut sister arriving at Alpha Centauri 10 years after she left, and sees his sister's clock register those 10 years. When Astrid arrives at Alpha Centauri she sees Terence's clock register that ten years have passed, and sees him aging a year per year as she returns home.

Now let's assume a Dopplerian universe. In this universe, it is known that Descartes was wrong not to have listened to his mentor Beeckman: light travels at a finite speed c (in Terence's Tellurian frame of reference). Otherwise the universe is classical, as before. Now things are interestingly different for the interstellar twins: because light takes 6 years to travel from Alpha Centauri to Earth, when Terence actually sees the event of Astrid's arrival there, 16 years will have passed since they parted! He sees Astrid's clock register only 10 years while his has registered 16. (Astrid's clock is apparently running slow by a factor 5/8 compared to his.) He is then even more surprised to see his sister take only 4 years to return, and watches his sister's clock running at 2.5 times the speed of his own;

[6] Cf. Lawrence Sklar (1974, 272): "One must always be careful in special relativity to distinguish what an observer actually *sees*, literally, from what he computes to be the case".

[7] Davies (1995, 59–65) considers the twin paradox by "equip[ping] our twins with a powerful telescope so that they can watch each other's clocks throughout the journey," and then discusses what clock readings they would see, distinguishing the Doppler effect from the time-dilation effect.

thus, as Terence views Astrid's return trip and all the processes happening in it, he sees them appearing to occur four times as fast ($2.5 \div 5/8$) as during the trip outwards! Astrid has an analogous experience. When she arrives at Alpha Centauri, she observes Terence's clock to be reading only 4 years. For the image of the clock registering 4 years travels the 6 light-years to Alpha Centauri to arrive there 10 years later (all from Terence's frame of reference). So Astrid sees Terence's clock has been running 2.5 times as slowly as hers (i.e. at 2/5 speed)! But on the way home she sees it to be running 1.6 times as fast: in the ten years it takes Astrid to return, Terence appears to her to age 16 years! Thus she, too, is puzzled to see her twin's clock going 4 times as fast ($1.6 \div 2/5$) as it was on the outgoing leg of the journey. But the twins' puzzlement is relieved when they learn about the Doppler effect: Events and processes occurring in a frame of reference in motion towards an observer appear to be speeded up ("blue shift"); occurring in a frame of reference in motion away from an observer they appear to be slowed down ("red shift").

Still, this does not explain the discrepancy in their experiences. Granted there is a certain symmetry: each twin sees the other's clock running 4 times as fast on the return trip as it was on the way to Alpha Centauri. But if all inertial motion is relative, they should have experienced the same red shift while moving apart, and the same blue shift when moving back towards one another. That they didn't is explained by the fact that we have taken the speed of light to be c in Terence's frame only. If the speed of light is also c in Astrid's frame of reference, then the situations are entirely symmetrical, and Astrid should see Terence's clock run slow on the outward trip by a factor of 1.6, and fast on the return journey by a factor of 2.5. Thus when she reaches Alpha Centauri, she should see Terence's clock read $10 \times 5/8 = 6.25$ years. But then Terence would age $20 - 6.25 = 13.75$ years while she returns. This is a long way short of his aging calculated by the Doppler factor 2.5, $10 \times 2.5 = 25$ years, an obvious impossibility! In short, the assumption that the speed of light is the same in all frames of reference is at variance with the assumption of classical physics that all inertial motion is relative.

This, then, is the kind of discrepancy that physicists faced at the beginning of the twentieth century. All the experimental evidence seemed to suggest that the speed of light is the same in all inertial reference frames. But this is incompatible with the requirement that everything will appear the same from each inertial reference frame, unless something else gives. In terms of the twin example, the only way for Astrid to see a Doppler effect equal to Terence's is if the length of the journey in a frame of reference in relative motion were to suffer a contraction along the direction of motion by a factor of $\sqrt{(1 - v^2/c^2)} = \sqrt{(1 - 0.36)} = 0.8$. (That the dimensions of a body are distorted in this way by their motion through the aether was independently suggested by George FitzGerald and Hendrik Lorentz,[8] taken up by Joseph Larmor, and generalized by Henri Poincaré.) But this also implies a similar effect on the rates at which processes occur: they must slow up by the same factor in a frame of reference in relative motion. In terms of the present example,

[8] Harvey Brown, in his (2005, 3, 45–55), explains that the conjecture of FitzGerald and Lorentz was that the dimensions of the body were altered in a certain ratio, not that there was a physical contraction along one of them. See also Mauro Donato, "Relativity theory between structural and dynamical explanations," forthcoming in *International Studies in Philosophy of Science*, preprint p. 7.

if the twin in motion at $0.6c$ covered a distance of only 4.8 light years (6×0.8) in her own frame of reference, then this would take her only $4.8 \div 0.6 = 8$ years ($= 10y \times 0.8$) in that frame. Time for the moving twin would, from the point of view of the stationary one, run more slowly by a factor of 0.8. This is the so-called *time dilation* effect, partially understood by Larmor and Lorentz, but first explicitly articulated by Einstein,[9] and now known to be a really occurring effect. Lorentz and company, of course, assumed that all these dilations and contractions could be referred to the frame in which the aether is at rest, and would effectively prevent one from detecting which frame this is. This is where Einstein stepped in: he dispensed with the unobservable-in-principle aether, so that each inertial frame would be on the same ontological footing.

So let us assume, finally, an Einsteinian universe for our twins. As before, when Terence actually sees the event of his sister's arrival at Alpha Centauri, his own clock registers 16 years. But, because of time dilation, Astrid's clock registers only 8 years, and thus *appears* to Terence to be running slow by a factor of 2, i.e. 1/2 as fast as his. On the return leg, Terence sees his sister's clock advance 8 years whilst his only advances 4; so the clock (and the aging processes of his sister and everything moving with her) *appear* to be running fast by a factor of 2. Despite these appearances, of course, Terence (who is now a whiz at physics) can *infer* that in each case the effect is 0.8 times what would be expected from the Doppler effect alone: a lag by a factor of 5/8, multiplied by 0.8, gives 1/2; a speeding up by a factor of 2.5 times 0.8, yields 2. Thus he *infers* that Astrid's clock is running slow because of time dilation.

Astrid, on the other hand, on looking back to Earth as she is arriving at Alpha Centauri 8 years later, sees the Tellurian clock register 4 years, as before. By Astrid's reckoning, the image has travelled eight years to get to Alpha Centauri, so Terence's clock *appears* to Astrid to be running slow by a factor of 2. On the return leg, Astrid sees her brother's clock advance its remaining 16 years whilst hers only advances 8; so the clock (and the aging processes of her brother and everything moving with him) appear to be running fast by a factor of 2. Again, she can calculate that since the effects should have been 5/8 and 2.5 if they were due to the Doppler effect alone, the difference is due to the fact that Terence's time is slowed relative to hers by the time dilation factor 0.8. (During her 8 years on the outward leg, she sees Terence's clock move 4 years when it should have moved 5 by the Doppler effect alone, since 8 times the Doppler effect of $5/8 = 5$; on the way back she sees Terence's clock move 16 years instead of the 20 that would be 8 times the Doppler effect of 2.5). Thus the situation is entirely symmetrical: while they are in relative motion, each twin suffers an *inferred* time dilation, a slowing-down of the aging process, from the point of view of the other. And in terms of what *appears*, they both see their twin sibling's clock running slow by a factor of 2 while they are moving apart, and running fast by a factor of 2 when they are approaching one another.

This shows us that the scenario depicted is entirely consistent. But how does it resolve the paradox? If everything is symmetrical, then why don't the twins

[9] For an authoritative discussion of the extent to which Larmor, Lorentz, and Poincaré did and did not anticipate Einstein's discovery of time dilation, see Brown (2005), esp. Ch. 4.

age by the same amount? The usual (and correct) explanation is as follows: although while the twins are in constant relative motion the situations are indeed perfectly symmetrical, this is not so for their journeys or paths through spacetime as a whole. For in this thought experiment, the terrestrial twin Terence undergoes no acceleration, whilst his adventurous astronaut sister must decelerate through $-0.6c$ on arrival at Alpha Centauri, and then accelerate through another $-0.6c$ to the same speed in the opposite direction. On the other hand, in the idealized conditions of the thought experiment, any time dilations due to the accelerations are ignored. But although this explanation is perfectly correct, it leaves a lingering sense of puzzlement. If the difference in the ages of the twins is not due to any time dilation caused by acceleration, and yet while they are in inertial motion relative to one another each sees the other's time dilated by the same factor, *how does the asymmetry in the paths taken result in a difference of time elapsed for each twin? How is a difference in paths even relevant to the situation*? To many people, this has suggested that the difference in the twins' ages *is* due to a time dilation caused by acceleration.

4. MODIFIED TWIN PARADOX

Thus it is often said that the reason for the discrepancy in the twins' ages is that whereas Terence is in inertial motion throughout, Astrid is the one who really moves because of her acceleration, although this acceleration lies outside the scope of the theory. This seems to imply that it is Astrid's (here instantaneous) non-inertial motion that is responsible for the dilation.[10] Some even go so far as to claim or imply that Special Relativity applies only to systems in inertial motion,[11] and that a proper resolution of the paradox must therefore involve General Relativity.[12] But this is incorrect. On the one hand, Special Relativity is perfectly applicable to accelerated motions, and on the other, although the fact of Astrid's acceleration is a necessary condition for their taking different paths through spacetime, the time dilation is due to the different paths, not an effect of the acceleration itself.[13]

[10] See for instance J.J.C. Smart, (1968, 231): "The clock paradox comes from the following fallacious bit of reasoning. In our calculations we have taken Jack to be at rest and Jim to be moving with a velocity of either $+v$ or $-v$ relative to him. Equally, it is said, we could take Jim to be at rest ... The fallacy in the reasoning is that the first calculation (showing Jim to be younger than Jack) was correct, because Jack has been in the same inertial system throughout. However Jim had to be accelerated at Alpha Centauri ..."

[11] To cite two contemporary examples from the world wide web: "However, this resulted in a limitation inherent in Special Relativity that it could only apply when reference frames were *inertial* in nature..." (http://en.wikipedia.org/wiki/Inertia); "This dilemma highlights a limitation of the Special Theory of Relativity that we have already alluded to. It only applies to observers in uniform motion, and not to accelerated frames." (http://theory.uwinnipeg.ca/mod_tech/node141.html).

[12] Cf. this analysis on the *Encyclopedia Britannica* internet site: "The answer is that the paradox is only apparent, for the situation is not appropriately treated by special relativity. To return to Earth, the spacecraft must change direction, which violates the condition of steady straight-line motion central to special relativity. A full treatment requires general relativity, which shows that there would be an asymmetrical change in time between the two sisters. Thus, the "paradox" does not cast doubt on how special relativity describes time, which has been confirmed by numerous experiments." http://qa.britannica.com/eb/article-252886.

[13] One suspects that Einstein himself unwittingly contributed to this misunderstanding by using arguments in his (1918) from General Relativity to defend the consistency of the Special Theory. But in fact what Einstein is defending is not the self-consistency of SR attacked by Dingler, but the consistency of the account of *time dilation due to accelerated motion* in SR with the General Relativistic equivalence of acceleration with gravity.

To see this, we can construct a journey for each of the twins with equivalent non-inertial paths as follows. Suppose the trusty Terence tires of his terrestrial tenure, and takes residence in Telstar, a nearby space-station, a doughnut shaped ship that simulates gravity by rotating. To relieve tedium, Terence sets it spinning very fast about an inertial point, so that for precisely the period (say, 2 years, in the terrestrial frame) in which Astrid is decelerating at $-0.6c$ a year, and undergoing a corresponding time dilation due to this acceleration (leaving her D years younger, where $0 < D < 2$), Terence undergoes an exactly corresponding time dilation due to his rotational acceleration. Now when Astrid returns, she will be between 16 years and 18 years older, and Terence will be between 20 and 22 years older, and they will differ in age by exactly 4 years. In the reference frame of the inertial point near earth, exactly 22 years will have passed.[14] But in their paths through spacetime, the twins will have been in inertial motion for the same 20 years with respect to that point and its inertial frame, and will have undergone time dilation due to their accelerations for the same two years in that frame. Yet one is still 4 years older than the other. It follows that it can't be said that Special Relativity applies only to systems in inertial motion—if that were so, there could be no explanation of the Twin Paradox in the theory. But we have just so explained it! Thus *the difference between the ages of the twins is not due to one's being in inertial motion, the other not*. Both their ages are true measures of time, in the original thought experiment, as well as in this modified version.

The correct conclusion is that it is not any difference between inertial frame-times that accounts for the difference in the twins' ages, but the difference in their paths through spacetime. It is the time elapsing along a particular path in spacetime that measures how fast the processes traversing that path are going, how fast the people or things undergoing them are aging, how fast they are becoming. In the non-Euclidean metric of Minkowski spacetime, it is the longest, not the shortest, time interval between two spacetime points that is given by the straight line in spacetime connecting them.[15] The longer the spacetime path between them, the shorter the time elapsed along that path. In the original thought experiment, Astrid travels along two sides of a triangle, and Terence by the remaining side; in the modified version, Astrid's straight lines are joined by a curve, while Terence's straight line is interrupted by a spiral of the same length. In each case Astrid's path is longer, and the time elapsed shorter. It follows that it can't be said, as one often reads, that the duration of processes in relativity theory is relative to an inertial frame. In the sense of time lapse that is relevant to the twin paradox—how much time elapses for each twin—*it is simply false that time lapse is frame-dependent*, i.e. depends on the inertial frame adopted. Indeed, the duration of each twin's journey through spacetime is an invariant measure: it is the same in all inertial reference frames.

[14] Of course, the physical situation could be made even more realistic, if desired, by having Astrid accelerate away from Earth to his speed of $0.6c$, and then decelerate to zero on return. But again, this could be compensated for by having Terence spin his Telstar for the same period.

[15] This is not to say that this feature of the Minkowski metric is *causally responsible* for the fact that the time lapse is longest along an inertial path, an interpretation against which Harvey Brown has argued in detail in his (2005).

5. PROPER TIME AND PROPER LENGTH

As I stated at the start of this essay, there are in fact two different measures of time in relativity theory: they have different formal measures, and different ontological baggage. This parallels the case for mass. In each case, what in classical physics had been thought to be a univocal or absolute property of the system turned out to be degenerate. For in the transition to Einstein–Minkowski physics mass bifurcates into the relativistically invariant *proper mass* m_0, and the *relative mass* μ, or mass in an inertial frame in which it is moving at a speed $v = \beta c$, whose quantity is a factor $\gamma = (1 - \beta^2)^{-1/2}$ times the proper mass.

But, as I have suggested above, I believe much of the confusion about relativity comes from interpreting proper time as if it is simply the relative time of an observer in her own rest frame. This misinterpretation is encouraged by the analogy with mass, but even more so, I will now suggest, by reading the case of time or *duration* as an exact analogue of that of space or *length*.[16] For the degeneracy of time in relativity theory is paralleled by a similar bifurcation in the concept of length. A body that is moving at a speed $v = \beta c$ with respect to a given inertial reference frame will, as already discussed, undergo a length contraction in the direction of its motion, so that its length $L = L_0/\gamma$, where L_0 is its *proper length*. The latter is defined as its length in the rest frame: if $v = 0$, $L = L_0$. Analogously, it may be thought, any periodic processes associated with he body will suffer a time dilation, so that $t = \gamma t_0$, with the result that in the rest frame where $v = 0$, $t = t_0$. *Proper time*, then, it may be asserted, *is just* t_0, *the time coordinate as measured in the rest frame*.

But this it is not! Proper time was introduced by Hermann Minkowski in his famous 1908 paper (Lorentz et al., 1923, 73–91) as follows. If at any point $P(x, y, z, t)$ in spacetime we imagine a worldline running through that point, the magnitude corresponding to the timelike vector dx, dy, dx, dt laid off along the line is

$$d\tau = \sqrt{\left(c^2\,dt^2 - dx^2 - dy^2 - dz^2\right)}/c$$

Proper time is now defined as the integral of this quantity along the worldline in question. Introducing the concept, Minkowski wrote: "The integral $\tau = \int d\tau$ of this quantity, taken along the worldline from any fixed starting point P_0 to the variable endpoint P, we call the proper time of the substantial point at P." (85) As he proceeded to explain, x, y, z and t—the components of the vector OP, where O is the origin—are considered as functions of the proper time τ, and the first derivative of the components of this vector with respect to the proper time, $dx/d\tau$, $dy/d\tau$, $dz/d\tau$ and $dt/d\tau$, are those of the *velocity vector* at P.

It is a consequence of this definition that the element of proper time $d\tau$ is not a complete differential. Arnold Sommerfeld, in his notes appended to Minkowski's 1908 paper when it was reprinted in a book (Lorentz et al., 1923, 92–96), remarked that Minkowski had mentioned this to him. He comments:

[16] I am indebted to Storrs McCall (private communication) for suggesting to me the relevance here of the analogy with proper length. I am also indebted to Kent Peacock for helping me eradicate some infelicities in my discussion of this in an earlier draft.

> [T]he element of proper time $d\tau$ is not a complete differential. Thus if we connect two world-points O and P by two different world-lines 1 and 2, then
>
> $$\int_1 d\tau \neq \int_2 d\tau$$
>
> If 1 runs parallel to the t-axis, so that the first transition in the chosen system of reference signifies rest, it is evident that
>
> $$\int_1 d\tau = t, \quad \int_2 d\tau < t$$
>
> On this depends the retardation of the moving clock compared with the clock at rest. (94)

Evidently, Sommerfeld had already resolved the twin (clock) paradox in 1923 in essentially the same terms as I have given above.

What is crucial to this resolution is that the proper time calculated along a given path in spacetime is an *invariant quantity*: it retains the same value under transformation of inertial frame. It is for this reason that it "can claim the prerogative of representing the objective lapse of time", to use Gödel's own words (558), and thus undermines his argument from the relativity of simultaneity to the unreality of time. Of course, Gödel assumed that an objective lapse would have to consist in a *global* plane of becoming, and therefore could not be relative to spacetime path; but, according to the point of view I am advocating here, this assumption is unwarranted in relativistic physics, where becoming is *local*, and dynamical change is parametrized by proper time, not co-ordinate time. It remains the case, of course, that the proper time is a maximum in the rest frame of an inertially moving object, and that in this circumstance it is numerically equal to the co-ordinate time. For when $v = 0$, $\beta = 0$, so that $\gamma = (1 - \beta^2)^{-1/2} = 1$, and $\tau = t_0$. But this is only numerical equality, not identity. It corresponds to the fact noted above that the longest time interval between two spacetime points in timelike separation is given by the straight line in spacetime connecting them. All other paths, whether the two inertial paths of the original Twin Paradox thought experiment, the paths of the Modified Twin Paradox incorporating a curve of deceleration and a spiral, or even a steady curve representing the travelling twin gradually slowing up turning round and returning, are shorter. But by the same token, Special Relativity is perfectly able to account for these non-inertial paths, and for each of them the proper time would be calculated by an integration along the path in question, not by the difference in time co-ordinates in any inertial frame. If proper time were the time co-ordinate in an inertial frame at rest, t_0, it would not be applicable to such curved paths. In contrast, proper length is the length of an object—a metre stick, say—in a specific frame of reference, namely, the inertial frame in which it is at rest.

Still, it may be objected, proper length is nevertheless also an invariant quantity. Just as the length of a path joining two events in timelike separation is invariant under change of frame, so is the length of a curve joining two events in spacelike separation. Indeed, it is often argued that the analogy between it and proper time is perfect: "proper length is the invariant interval of a spacelike path

whereas proper time is the invariant interval of a timelike path".[17] Thus, it is suggested, the definition of proper length should be generalized so that it is the exact analogue of proper time: a line integral along a curve joining two spacelike separated events. But a little further reflection shows that this cannot be right: an arbitrary curve joining two spacelike separated events is not generally a length. It is only a length if all the points on the curve are simultaneous in some given reference frame. And while the path integral along such a curve is indeed independent of the choice of reference frame, it has no particular physical significance. It does not even represent a *path*, in the normal acceptation of a path as a series of positions that can be successively traversed—as, for instance, by Harvey Brown's waywiser (Brown, 2005, front cover, p. 8)—for such an interval is timelike. Proper length is correctly defined as the path integral, not along *an arbitrary curve* joining the endpoints of the metre stick at the same time, but along *the shortest curve*, which is a straight line joining them in the frame at which they are at rest. If (elapsed) proper time were the strict analogue of this, it would be the longest time between two timelike separated events, which would be the time in a frame of reference at rest, i.e. the co-ordinate time. It is precisely this interpretation that I am attacking. Because proper length is the interval between two events *at the same co-ordinate time*, it is specific to a particular reference frame.

Thus *proper time* has a fundamentally different character from proper length. Although both are invariant under change of frame, proper length is the length of an object in its own rest frame, whereas proper time is independent of frame. In this respect proper length is analogous to proper mass. (It differs from the latter, however, in that proper mass seems to be an essential characteristic of an elementary body (such as an electron), whereas proper length is a contingent one.) At any rate, there is a fundamental dissymmetry between duration and length in Special Relativity, somewhat obscured by talk of their embodiments in observers' clocks and rods. For whereas an observer's clock measures proper time elapsed along a path, a dynamical variable specifiable independently of reference frame, the proper length of the observer's measuring rod is specific to the inertial frame in which the observer is at rest. Thus we see that, ironically, there is a sense in which Minkowski's introduction of proper time undermines his famous pronouncement at the beginning of his paper about the demise of time: "Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality". (75)[18]

Interestingly, this very lack of a frame-independent invariant length in relativity theory leads to another paradox, although this paradox depends on a combination of quantum theory and relativity theory. Even though it is something of a digression from my main argument, it seems to further accentuate the disanalogy between space and time. The point is this: in the attempt to come up with

[17] Quoted from an article on proper length in Wikipedia (http://en.wikipedia.org/wiki/Proper_length: May 5, 2007). The author suggests a generalization of proper length so that it is given by the line integral $L = c \int_P \sqrt{[-g_{\mu\nu}\, dx^\mu\, dx^\nu]}$, where $g_{\mu\nu}$ is the metric tensor for the spacetime with $+\ -\ -\ -$ signature, normalized to return a time.

[18] Minkowski's judgement is echoed by Einstein in his essay "The Problem of Space, Ether and the Field in Physics": "Hitherto it had been silently assumed that the four-dimensional continuum of events could be split up into time and space in an objective manner... With the discovery of the relativity of simultaneity, space and time were merged in a single continuum..." (1954, 281–282).

a theory of quantum gravity, all researchers are agreed on the fundamentality of the Planck length, $L_P = \sqrt{\hbar G/c^3}$, where $\hbar$ is Dirac's constant (Planck's constant divided by 2π), and G is the gravitational constant: this is extremely small, about 1.6×10^{-35} metres. In some sense this should represent the smallest length there is, and be so for all possible observers. And this is where the paradox gains purchase: according to Special Relativity, there is no frame-independent invariant length. If the Planck length L_P is regarded from a frame of reference moving with respect to it at $v = \beta c$ it will, as already discussed, be contracted in the direction of its motion, so that its length will be less, namely $L = L_P/\gamma$. But by hypothesis, nothing is shorter than the Planck length. In other words, Einstein's relativity theory—which in any case may be expected to break down at this scale—is in contradiction with the posited invariance of the Planck length, since in it there is no preferred reference frame. This paradox was proposed by Giovanni Amelino-Camelia in 1999, who at the same time proposed a solution to it with the so-called theory of Doubly Special Relativity (or DSR).[19] The principal idea of this theory is that there is not just one invariant constant, the speed of light c—which is now the speed of very low energy photons—but also the Planck length. If this theory is true, of course, the invariance of the Planck length is analogous to that of c, not to that of the proper time, since it is a universal constant, whereas proper time is a variable. But whatever the fate of DSR, it seems that time and space are on very different footings in relativity theory.

I am by no means the first to point out the radical implications of relativity theory for our understanding of time. As Milič Čapek has stressed in several publications (1966, 1975, 1976), the invariance of Minkowski's relations of being in the absolute past or future of an event means that in relativity theory the role of time is strengthened and made more distinct than in classical physics. The distinction between proper time and coordinate time is stressed by Larry Sklar in his treatment of the clock paradox;[20] and Kent Peacock (1992), has also discussed the paradox in terms of a comparison between the proper times of the twins while they are spatially distant. But perhaps the clearest explanation of the distinction between "time co-ordinate" and "proper time" and its significance was given by Howard Stein:

> Proper time is not a quantity attached to space-time points or to pairs of space-time points; it is in this respect a notion utterly different from the quantity "time" or "time interval" of pre-relativistic theory… The fundamental physical role of proper time comes from the principle (here stated roughly) that whenever a process takes place along a well-defined line of

[19] Giovanni Amelino-Camelia (2001); see Smolin (2006, 365). H.S. Snyder had already resolved the incompatibility between the existence of a minimum length and the requirement of Lorentz invariance in his (1947), positing two invariant scales and a non-linear basis for the Poincaré algebra. See also Joy Christian (2004), whose theory posits only one observer-independent conversion factor, the inverse of the Planck time, with c emerging as an invariant but derivative factor.

[20] Sklar (1974, 268) correctly points out that, whereas " 'co-ordinate time' between two events is relative to a given inertial frame", "[p]roper time is defined only for events at timelike separations and only relative to a particular spacetime curve between the events. On the other hand it is an invariant notion." Unfortunately, though, he goes on to claim that anyone who wishes to assert that future events are not real relative to the assertor is forced by the Putnam–Rietdijk argument to admit that such notions are "just as relative to an inertial state of motion of the assertor and just as 'nontransitive across observers in different states of motion' as we have made the simultaneity relation." (275).

> space-time ("world-line"), the time rates in the dynamical principles that govern that process are to be understood in terms of the proper time along that line (and *not* in terms of a "time coordinate"...).[21]

Yet it seems to me that the significance of this degeneracy of time in relativity theory is still largely unrecognized. Philosophers and physicists continue to write as if it is the time co-ordinate function, or time in relation to an inertial observer, and not proper time, that measures the duration of processes in relativistic physics. This is implicit in all discussions that agree with Gödel in construing the objective lapse of time in terms of an infinity of layers of "now", with these planes of simultaneity picked out by the time co-ordinate function in an inertial reference frame, such as the arguments of Putnam, Rietdijk and Maxwell discussed above.

6. CONCLUSION: PROCESS AND BECOMING

Finally, I want to return to Gödel's attack on the reality of time based on his argument from General Relativity. It is not generally recognized that this same conception of time lapse in terms of the planes of simultaneity parametrized by a time coordinate function underlies this argument too.

Having stated his objections to the relativity of time lapse, Gödel considers a way of avoiding this relativity first proposed by the astronomer James Jeans in 1935. This is founded on the observation that the expansion of the universe does after all allow (in principle) the singling out of one set of preferred reference frames, relative to which a notion of cosmic time can be defined, namely the family of frames tracking the mean motion of matter. As Gödel reports, Jeans had proposed this as a way of recovering "the intuitive idea of an absolute time lapsing objectively".[22] At any rate, when we talk about the Big Bang having occurred 15 billion years ago, it is with respect to such a cosmic time function that the age of the universe is being calculated.

It is against this Jeansian scenario that Gödel directs his argument from General Relativity. If the incongruity of having time lapsing at different rates for different observers is to be circumvented by an appeal to a cosmic time function, what if there are scenarios where no such cosmic time function can be constructed? This motivates his construction of solutions to Einstein's field equations representing static, spatially homogeneous universes, rotating with respect to the totality of galactic systems. In these "rotating universe" solutions, such "an *absolute* time

[21] Stein (1968, 11, fn. 6). This quotation from Stein was my starting point in the line of argument for his paper. Cf. also p. 16: "...*'a time co-ordinate'* is not *'time.'* Neither *a* nor *b* is, in any physically significant sense, 'present' (or past) for any observer at *c*—regardless of his velocity—for neither *has already become* for *c* (nor has *c* for them); but *a* has already become for *b*, and can influence it." [Here *a* and *b* are connectible by a time-like vector *ab*, the other pairs by space-like vectors *ac* and *bc*.]

[22] Gödel (1949b, 559); James Jeans (1935, 22–23). As Gödel observes, the "mean motion of matter" is not a very precise notion, and is contingent on facts about our cosmos: "What may be called the 'true mean motion' is obtained by taking regions so large that a further increase in their size does not any longer change essentially the value obtained. In our world this is the case for regions including many galactic systems" (1949b, 559, n. 7). This approximation could perhaps be improved, but would still involve "introducing more or less arbitrary elements (such as, e.g., the size of the regions or the weight function to be used in the computation of the mean motion of matter)" (560, n. 9).

does not exist" (Gödel, 1949a, 447): that is, because of the existence of closed time-like curves ("time loops") through every spacetime point of the solution, "there can be no cosmic time coordinate t in M which increases along every future-directed time-like or null curve" (Hawking and Ellis, 1973, 170). The *coup de grace*—"strengthening further the idealistic viewpoint"—is then provided by the fact that the existence of these time loops implies that in Gödel's rotating universes it would be possible (in principle) "by making a round trip on a rocket ship in a sufficiently wide curve" for someone "to travel into any region of the past present or future", and in particular "into the near past of those places where he has himself lived" (Gödel, 1949b, 560–561). This opens up the whole Pandora's box of paradoxes concerning killing one's own parents or grandparents and thus preventing one's own conception, towards which much philosophical attention has been directed.

Since Gödelian universes are causally pathological, physicists have preferred to exclude their possibility by stipulation. Thus Hawking and Ellis lay down as a postulate "that space-time satisfies what we shall call the *chronology condition*: namely, that there are no closed timelike curves" (1973, 189). An apparently stronger condition is the *causality condition*, which holds if there are no closed time-like or null geodesic curves, although Hawking and Ellis give an argument to prove that "in physically realistic solutions, the causality and chronology conditions are equivalent" (192). Stronger than this is the *strong causality principle*, which holds at some point p "if every neighbourhood of p contains a neighbourhood of p which no non-spacelike curve intersects more than once" (192). Any of these conditions, laid down as a condition for the physicality of a universe (and thus as a constraint on the viability of solutions to Einstein's field equations), will preclude Gödel's cylindrical universes by *fiat*. But it does not seem likely that Gödel would have found such attempts to parry his arguments any more compelling than the Jeansian strategy he was refuting. His intuition is essentially Kantian: if time lapse is objective, it must be a feature of any possible universe, including those that do not expand. The fact that ours is an expanding universe is a contingent fact—at least, according to our current understanding: there is no known way of deriving the fact of the expansion of the universe from physical law. Likewise, if there is no way of deriving the causality condition from first principles, then, as a rebuttal of Gödel, it has no more force than the brute empirical fact of the expansion of the universe. An *a posteriori* axiom does not have the requisite power to establish what must be true *a priori* of any possible universe. If it is possible in principle for there to be a universe to which time lapse is not applicable, then this is enough, by Gödel's lights, to refute the objectivity of time lapse.

But does Gödel's argument prove that time lapse is self-contradictory? The scenario he depicts is of two space travellers traversing different paths through his cylindrical spacetime—one from P to Q (perhaps an indefinitely small distance away) along a worldline, one from Q in a timelike curve back to P—in such a way that when they meet back up their clocks do not agree on how much time has elapsed. Gödel argues that if there are 3-spaces "which are everywhere spacelike and intersect each worldline of matter in one point", then "time measured along the world lines of matter in their positive direction would yield a coordinate system with the property that the 0th coordinate always increases if one moves in a

positive time-like direction" (Gödel, 1949a, 447, 449), in contradiction to the existence of the time loop scenario described, "which implies that all coordinates of the initial and the endpoint of a time-like line [e.g. here any of the timelike lines from *P* through *Q* back to *P*] are equal in certain cases" (449). That is, if time lapse is to be measured by a (any) time co-ordinate function, then, despite the time taken for the traveller's trip, it will come out to be 0 according to that time-coordinate function.

But if this proves time lapse unreal, then this is already proven in Special Relativity, by the case of the Twin Paradox considered above! The problem with Gödel's formulation, as I hope should be clear by this juncture, is that it fails to appreciate the degeneracy of time: time lapse is not represented in relativity theory by the time co-ordinate function, but by proper time.[23] This is the time as measured for the paths through spacetime, and, as we have seen, these will not in general be the same in any spacetime where the motions along those paths involve asymmetric accelerations. The time elapsed for the traveller traversing a time loop, as measured by the proper time, will generally be quite considerable, whereas for a traveller who has travelled from *P* to a point *Q* an arbitrarily small spacetime interval away, the proper time elapsed will be arbitrarily close to zero. Thus, as paradoxical as the scenario depicted by Gödel is, it does not refute the possibility of time lapse. It precludes time lapse as he conceived it, in terms of global planes of simultaneous becoming; but not time lapse conceived, as I have urged is implicit in relativity theory, as the unfolding of processes along worldlines. What, then, are the implications of Gödel's scenario for our understanding of time?

I think it is crucial to remember here that spacetime in general relativity theory is not a background, nor is it a perduring space which we may imagine a traveller travelling through and exploring. One cannot *have* the spacetime and *then* superpose trajectories or events onto it: this is to confuse spacetime with a perduring background space. Any process or events must be represented in the spacetime. So in this case it is legitimate to ask: does a worldline looping from *P* through *Q* and back to *P* represent a possible process? Notice that this is to pose a question that is slightly different from the usual one about the possibility of confronting one's younger self: I am asking, is it possible for a continuous worldline to loop from *P* back to the same event *P*? Could one have a space traveller, or any other object large enough to bear traces of aging, both bear and not bear those traces at *P*? Clearly not.[24]

Now, since aging entails an irreversible process, this might seem to suggest that perhaps a simple reversible process *could* lie on such a loop. I would argue, however, that not even this is possible: an individual process (and here we are idealizing a process as 1-dimensional) must be from some one point-event *a* to another *b*, and cannot also be from *b* to *a*, without violating the idea of what a process is. That is, the relation of becoming—one individual event *x*'s coming out

[23] The same point has been stressed by Dennis Dieks in his (2006): "The rate of these local processes is determined by the amount of *proper time* between events, and not by differences in cosmic time" (167), see Arthur (1982).

[24] Commenting on such scenarios as travelling such a loop and preventing oneself from setting out in the first place, Hawking and Ellis comment: "Of course there is a contradiction only if one assumes a simple notion of free will" (1973, 189). That this is incorrect is shown by my example of an object traversing a time loop through some point *P*, which would have to both bear and not bear traces of aging at *P*.

of another y, xBy—is intrinsically asymmetric: $(\forall x, y)(xBy \rightarrow \neg yBx)$.[25] Now, of course, transitivity might fail: one might have local becoming along "sufficiently small" segments of a worldline, a piece-wise, fragmentary process of local becoming along individual worldlines, not adding up to "global becoming along the loop".[26] But such a non-additivity of processes seems (to me at least) to be as contrary to our intuition of process as a failure of asymmetry would be. If this is so, then in the spacetimes that are solutions to Einstein's field equations, this requirement of the asymmetry and transitivity of process on lines that represent processes entails the chronology condition of Hawking and Ellis: there must be no closed timelike curves.

Now if this analysis is correct, I believe we have a full answer to Gödel. He is right to demand that if time lapse is to be counted as objective, it should both be invariant and be a feature of any possible universe. We have seen that the objectivity of time lapse is guaranteed by its being measured by the proper time, with proper time invariant under change of inertial frame. But in addition, in order for a timelike curve to represent the trajectory of a process, it cannot, because of the asymmetry and transitivity of becoming, be closed. But this means we have an argument for the chronology condition: it is an *a priori* condition for representing a possible process in spacetime.

This perspective seems to have implications for modern attempts to eliminate time from physics. Time, it entails, is not just a co-ordinate, appearing in the metric on a par with the three dimensions of space except for the factor of i (Stephen Hawking's "imaginary time": see his 1988, 134–139). It is fundamentally dynamical, tracking the evolution of all processes in the universe, and so cannot be rolled over into a space coordinate as we trace it backwards towards the universe's origins. The causal/chronological structure of timelike lines and worldlines are not symmetrical with spacelike lines in such a way as to be transformable away. In several approaches to quantum gravity, as Lee Smolin has argued, "*causality itself is fundamental*—and is thus meaningful even at a level where the notion of space has disappeared" (Smolin, 2006, 241). To this we may add: if it is indeed the case that time lapse is a necessary condition for process, then a world without time is a world without process. Now one is not obliged to depict interactions in particle physics, say, using the space-time representation in quantum physics: sometimes it is more convenient to use the energy-momentum representation. But to conclude from this that there is no time is to say that there is no process, and one wonders then what it is that is being represented. Similar misgivings, it would seem, should apply to Julian Barbour's claim there is no time in the most fundamental description of reality, the timeless universe he has dubbed "Platonia" (1999).[27]

To conclude: I have argued that time is degenerate in relativity theory. Coordinate time is used to track the synchrony of distant events, but it no longer

[25] This is a different question than that of the reversibility of processes, which concerns whether *types* of processes of a given kind always occur—whether nomologically or *de facto*—one way round with respect to an already given time-direction.

[26] This was suggested by the anonymous referee of this paper.

[27] Lee Smolin (2000) has given a detailed analysis of Barbour's argument, arguing that its conclusion can be resisted only if one restricts quantum cosmology "to theories in which all observables are accessible to real observers inside the universe" (23), and investigating in detail what this entails for cosmology.

has the classical role of tracking a worldwide hyperplane of becoming, as it did in classical theory. Instead it is proper time that measures time elapsed, and thus gives the true measure of the duration and rate of processes in spacetime. This bifurcation of time's roles is masked by a tendency to assimilate proper time to time in an observer's rest frame, by analogy with proper length, a tendency which is encouraged by unwarranted talk of an observer "inhabiting an inertial frame" and "experiencing" the events which are simultaneous with his or her state of consciousness. But whereas proper length is specific to a rest frame, proper time is not; its intervals are path-dependent, frame-independent, and invariant under change of reference frame. It is this proper time that measures the time elapsed for travellers in spacetime, which consideration is sufficient to resolve the Twin Paradox, as is shown with particular attention to what the twins could actually observe and infer about each other's times. It also disposes of Gödel's arguments for the ideality of time from the existence of closed timelike curves in certain General Relativistic spacetimes. On the contrary, it is argued, the condition that every timelike curve represent a possible process yields a justification for the chronology condition: because of the intrinsic asymmetry and transitivity of becoming, a closed time like curve cannot represent the path of a possible process.

ACKNOWLEDGEMENTS

I am indebted for their comments on earlier drafts of this chapter to Jim Brown, Vesselin Petkov, Kent Peacock, Storrs McCall, John Norton, and audiences in the Second International Conference on the Ontology of Spacetime and in the TaU Workshop following it (Concordia University, Montréal, 2006); although, of course, this does not imply their concurrence with the chapter's arguments. I am also much indebted to the anonymous referee for this volume for helpful criticisms of the penultimate draft.

REFERENCES

Amelino-Camelia, G., 2001. Testable scenario for relativity with minimum length. Phys Lett. B 510. arXiv: hep-th/0012238.

Arthur, R.T.W., 1982. Exacting a philosophy of becoming from modern physics. Pacific Philosophical Quarterly 63 (2), 101–110. April.

Arthur, R.T.W., 2006. Minkowski spacetime and the dimensions of the present. In: Dieks, D. (Ed.), The Ontology of Spacetime. Elsevier, Amsterdam, pp. 129–155. Chapter 7.

Barbour, J., 1999. The End of Time. Oxford University Press, Oxford, New York.

Brown, H., 2005. Physical Relativity: Spacetime Structure from a Dynamical Perspective. Oxford University Press, Oxford.

Čapek, M., 1966. Time in relativity theory: Arguments for a philosophy of becoming. In: Fraser, J.T. (Ed.), Voices of Time. Brazilier, New York, pp. 434–454.

Čapek, M., 1975. Relativity and the status of becoming. Foundations of Physics 5 (4), 607–617.

Čapek, M. (Ed.), 1976. Concepts of Space and Time. D. Reidel, Boston.

Christian, J., 2004. Passage of time in a Planck scale rooted local inertial structure. Int. J. Mod. Phys. D 13, 1037–1071.

Clifton, R., Hogarth, M., 1995. The definability of objective becoming in Minkowski spacetime. Synthese 103, 355–387.
Davies, P.C.W., 1989. The Physics of Time Asymmetry. University of California Press.
Davies, P., 1995. About Time. Touchstone, New York.
Dieks, D., 2006. Becoming, relativity and locality. In: Dieks, D. (Ed.), The Ontology of Spacetime. Elsevier, Amsterdam, pp. 157–176.
Dingle, H., 1972. Science at the Crossroads. Martin, Brian and O'Keefe, London.
Dorato, M., 2006. The irrelevance of the presentist/eternalist debate for the ontology of Minkowski spacetime. In: Dieks, D. (Ed.), The Ontology of Spacetime. Elsevier, Amsterdam, pp. 93–109.
Einstein, A., 1918. Dialog über Einwände gegen die Relativitätstheorie. Die Naturwissenschaften 48, 697–702. 29 November. An English translation exists on the web at http://en.wikisource.org/wiki/Dialog_about_objections_against_the_theory_of_relativity.
Einstein, A., 1954. Ideas and Opinions. Crown Publishers, New York.
Gödel, K., 1949a. An example of a new type of cosmological solutions of Einstein's field equations of gravitation. Reviews of Modern Physics 21 (3), 447–450. July.
Gödel, K., 1949b. A remark about the relationship between relativity theory and idealistic philosophy. In: Schilpp, P.A. (Ed.), Albert Einstein: Philosopher-Scientist. Tudor, New York, pp. 557–562.
Grünbaum, A., 1976. The exclusion of becoming from the physical world. In: Čapek, M. (Ed.), Concepts of Space and Time. D. Reidel, Boston, pp. 471–499.
Hawking, S.W., 1988. A Brief History of Time. Bantam Press, London.
Hawking, S.W., Ellis, G.F.R., 1973. The Large Scale Structure of Space-Time. Cambridge University Press, Cambridge.
Jeans, J., 1935. Man and the Universe. Sir Halley Stewart Lectures.
Lorentz, H.A., Einstein, A., Minkowski, H., Weyl, H., 1923. The Principle of Relativity. Methuen. Reprinted: Dover, 1952.
Maxwell, N., 1985. Are probabilism and special relativity incompatible? Philosophy of Science 52, 23–43.
Peacock, K., 1992. A new look at simultaneity. In: Hull, D., Forbes, M., Okruhlik, K. (Eds.), Philosophy of Science Association 1992, vol. I. Philosophy of Science Association, East Lansing, pp. 542–552.
Putnam, H., 1967. Time and physical geometry. Journal of Philosophy 64 (8), 240–247. April 27.
Rietdijk, C.W., 1966. A rigorous proof of determinism derived from the special theory of relativity. Philosophy of Science 33 (4), 341–344. December.
Savitt, S., 2006. Presentism and eternalism in perspective. In: Dieks, D. (Ed.), The Ontology of Spacetime. Elsevier, Amsterdam, pp. 111–127.
Sklar, L., 1974. Space, Time, and Spacetime. University of California Press, Berkeley.
Smart, J.J.C., 1968. Between Science and Philosophy. Random House, New York.
Smart, J.J.C., 1980. Time and Becoming. In: van Inwagen, P. (Ed.), Time and Cause. Reidel, Dordrecht.
Smolin, L., 2000. The present moment in quantum cosmology: Challenges to the argument for the elimination of time. Preprint, pp. 1–29.
Smolin, L., 2006. The Trouble with Physics. Houghton-Mifflin, Boston, New York.
Snyder, H.S., 1947. Quantized space-time. Physical Review 71, 38–41.
Stein, H., 1968. On Einstein–Minkowski space-time. Journal of Philosophy 65, 5–23.
Yourgrau, P., 1991. The Disappearance of Time: Kurt Gödel and the Idealistic Tradition in Philosophy. Cambridge University Press, Cambridge.

CHAPTER 12

On Temporal Becoming, Relativity, and Quantum Mechanics

Tomasz Bigaj*

Abstract

In the first section of the chapter, I scrutinize Howard Stein's 1991 definition of a transitive becoming relation that is Lorentz invariant. I argue first that Stein's analysis gives few clues regarding the required characteristics of the relation complementary to his becoming—i.e. the relation of indefiniteness. It turns out that this relation cannot satisfy the condition of transitivity, and this fact can force us to reconsider the transitivity requirement as applied to the relation of becoming. I argue that the relation of becoming need not be transitive, as long as it satisfies the weaker condition of "cumulativity": for a given observer the area of the events that have become real should not diminish as time progresses. I show that there are actually two relations of becoming that meet this weakened condition: Stein's (transitive) relation of causal past connectibility and the (non-transitive) relation that is the logical complement of the future causal connectibility. In the second part of the chapter I defend Stein's notion of temporal becoming against the attack that appeals to quantum-mechanical non-locality. I critically evaluate the argument given by Mauro Dorato (1996) that purports to show that space-like separated measurements done on the EPR system have to be mutually determinate. Finally, in order to account for the truth of counterfactual statements that link the space-like separated outcomes, I propose a dynamic conception of becoming, according to which the sphere of determinate events as of a given point may depend on the physical phenomena transpiring at this point.

1.

The special theory of relativity does not seem to be particularly conducive to the idea of objective temporal becoming. It is commonly held that relativity favors the "block universe" view (known also as "eternalism"), according to which all

* Institute of Philosophy, Warsaw University, Poland

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00012-0

© Elsevier BV
All rights reserved

events enjoy the same ontological status regardless of their location, and seriously undermines the views on time that distinguish spheres of a "determinate past" and an "open future". However, a dissenting position has been developed and defended by H. Stein in (1991) as a direct response to a paper by N. Maxwell (1985) and to earlier papers by H. Putnam (1967) and C.W. Rietdijk (1966). In his paper Stein shows how to define a notion of "real becoming" that is compatible with the postulate of relativistic invariance.[1] In order to achieve this, he focuses on the task of defining a binary relation between space-time points "being definite as of," which in his opinion can express philosophical intuitions regarding the division of the history of the world into an ontologically fixed past and a still open future.[2] Putting some reasonable constraints on such a relation, he then proceeds to prove the uniqueness of this relation, which basically reduces to the relation of causal connectibility. In what follows I argue that Stein's analysis is in one respect incomplete, as it leaves the complementary relation "being open (unsettled) as of" undefined. This is an important oversight, for it turns out that the relation of "openness" cannot possibly satisfy all the requirements imposed by Stein, and when we relax those requirements, we may end up with more than one possible notion of objective becoming that is relativistically invariant, contrary to Stein's uniqueness theorem.

Let us first present Stein's basic requirements regarding his definiteness relation. Let "Rab" be read as "the state at b is definite as of a" (or, alternatively, "for a, b has already become"), where a and b are space-time points. Stein insists that R should satisfy the following requirements in order to be a candidate for an expression of a notion of objective becoming:

(i) R is reflexive,
(ii) R is transitive,
(iii) For any point a there is a point b such that $\sim Rab$,
(iv) If b lies in the causal past of a (the past light cone of a), then Rab,[3]
(v) R is definable in terms of the Minkowski geometry of space-time.

Stein then produces a straightforward proof of the fact that the only relation satisfying all the conditions (i)–(v) is the relation of being in the causal past: Rab iff b lies in the causal past of a. In other words, the only reasonable candidate for the definite past as of a is the interior plus the surface of a's past light cone.

It may be instructive to see why another potentially feasible candidate—the outside of the future light cone of a—cannot play the role of the determinate past

[1] It has to be noted that Stein's proposal is not the first attempt to define temporal notions such as becoming, coming-to-be, or passing away in a relativistically invariant way. One particular proposal on how to equip these notions with a relativistically acceptable meaning has been put forth by Z. Augustynek in (1991), especially Ch. 6 "Past, present, future and becoming", which is a translation of his Polish book published in 1979. An interesting feature of Augustynek's conception is that he does not restrict himself to construing the aforementioned notions as relations between space-time points or events. Rather, he considers becoming, coming-to-be and passing away with respect to things (understood as four-dimensional manifolds of events).

[2] I leave open the question whether the proponents of the idea of the objective passage of time would be completely satisfied with Stein's conceptualization which analyzes the objective becoming as being relative to a given point. In particular, it may be pointed out that Stein's relation fails to do justice to our intuition of one, absolute "now" that is in constant flow. For a more detailed discussion on this subject see Savitt (2006) and the literature cited therein.

[3] R. Clifton and M. Hogarth (1995) have noted that this requirement can be replaced by the weaker condition stating that if b lies in the chronological past of a (i.e. b is located inside a's past light cone excluding its surface), then Rab. As a result, we would obtain a slightly different notion of becoming which is reducible to past *chronological* connectibility.

for a under Stein's provisions. Obviously the relation Rab, defined as "b lies outside the future light cone of a or $b = a$," is relativistically invariant, but it does not satisfy condition (ii)—it is not transitive. Clearly, when we select points b and c such that b is space-like separated from a and c is space-like separated from b, and c lies in the absolute future of a, b has to be judged as definite as of a, and similarly c as definite as of b, but c is not determinate as of a. And, as Stein has shown, if we took the transitive closure of R as the required relation, then it would become the universal relation, connecting every pair of points, thus clearly violating requirement (iii).

Stein stresses that the notion of the already definite part of space-time is relativized not to a temporal moment (as in classical physics) but to the "here and now" (space-time point). This is a straightforward consequence of the special theory of relativity which does not admit the notion of a spatially extended present.[4] Stein attributes the failure of the existing arguments against the definability of objective becoming in STR in part to the implicit assumption that there exists a set of present and spatially distant "actualities" (1991, p. 152). However, this does not mean that Stein's own notion of becoming cannot be applied to spatially extended objects. As W. Myrvold has recently shown (2003), an extension of Stein's definition of the definiteness relation R to arbitrary space-time regions can be achieved relatively easily. If α and β denote arbitrary space-time regions, then β is definite as of α iff for every space-time point $y \in \beta$ there is an $x \in \alpha$ such that Rxy. Of course one of the consequences of this approach is that there may be spatiotemporally extended objects such that they are neither fully definite nor fully indefinite as of α. Such a situation obtains when for some $y \in \beta$ there is $x \in \alpha$ such that Rxy, but for some $z \in \beta$ there is no $x \in \alpha$ such that Rxz.

My claim is, however, that Stein's analysis of objective becoming in Minkowski space-time has to be deemed incomplete, for a rather straightforward reason. The notion of temporal becoming requires distinguishing two spheres of space-time: a fixed sphere of things that have already become, and an open sphere of future possibilities. While Stein's relation corresponds to the first of these two spheres, we have to explicitly define another relation, circumscribing the sphere of future alternatives. In other words, we need a formal characteristic of a complement of the definiteness relation R, which can be referred to as "the indefiniteness relation S". Sab can be read as "the state at b is indefinite (open, unsettled) as of a". By analogy to Stein's procedure, we can stipulate in advance the requirements that the relation S should meet. Clearly, Stein's condition (i) should be replaced with its direct negation. S has to be irreflexive, because we don't want to admit that an event could ever be judged indefinite with respect to itself. However, regarding

[4] Some critics of Stein's conception focus their attack on this particular feature of his relation of becoming. For example, J. Faye, U. Scheffler and M. Urchs argue that one undesirable consequence of Stein's approach is the fact that two observers located at space-like separated points a and a' will disagree as to what events are determinate (1997, p. 36). Consequently, by relocating an observer in space we can make some events fall out of the realm of objective becoming. But, as Clifton and Hogarth have pointed out (1995, p. 383), even in Newtonian mechanics what is real for an observer depends upon his location in time. As space-time is a proper relativistic generalization of Newtonian time, we shouldn't be surprised that by transporting the observer from one location in space-time to another we can change his area of determinate events. I should also add that attempts to reconcile the requirements of the special theory of relativity with the existence of a spatially extended present have been made. For one of the most recent attempts along this line, which makes use of the inflationary Big Bang theory to pick a frame of reference that defines the absolute passage of time, see P. Forrest's "Relativity, the Passage of Time, and the Cosmic Clock" in this volume.

conditions (ii) and (iii) I don't see any *a priori* reasons why they should not be upheld in the case of *S*. It seems that if an event happening at *b* is not yet decided as of *a*, then whatever is not decided as of the undecided *b* should be treated—even more so—as undecided from the perspective of *a*. At least it can be claimed that the case for the transitivity of the relation *S* is as strong as the analogous case for *R*.

I have to admit, however, that this contention can be questioned by way of the following argument.[5] One may point out that the relation of indefiniteness *S* is usually characterized as the logical complement of the relation of being definite *R*, and, as a consequence, the requirement of the transitivity of the former seems to be ill-conceived. After all, there are convincing examples of non-transitive relations which complement transitive ones. For instance, the relation "*x* is part of *y*" is transitive, and yet its complement "not being part of" is clearly not transitive: when *x* is not part of *y*, and *y* is not part of *z*, *x* may still be part of *z*. By the same token, it can be claimed that while *R* should be transitive, its complement *S* does not have to be.

I have two replies to this objection. First of all, it ignores the fact that the relation between both notions *R* and *S* is entirely symmetric: if *S* is the complement of *R*, *R* is the complement of *S*. We should not be biased against the relation *S* by the fact that in natural language we refer to it using the term "indefiniteness" which contains the negative prefix "in". After all, we could select another term, for instance "openness", and insist that the relation *R* is defined as its negation "being un-open". So the mere fact that *S* is the complement of *R* does not give us strong enough reasons for the claim that it is *R* and not *S* which should be transitive. There have to be independent reasons for this claim (some of which will be briefly reviewed at the end of this part of the paper), and it seems to me that any prima facie argument in favor of the transitivity of the definiteness relation *R* can be modified as to support the transitivity of the openness relation *S*.

My second point is that so far I have not declared that *S* is to be defined as the logical complement of *R*. My intention is to treat the relation of indefiniteness as an autonomous ontological entity which expresses the commonsense belief that some events have the status of being "open" as of a given point. There may be some independent reasons (of which we will soon talk) why the relations *R* and *S* should be jointly exhaustive, i.e. for any two events *a* and *b* either *Rab* or *Sab* should hold, but this need not enter the definition of *S*. This observation should reinforce my claim that both relations *R* and *S* are on an equal ontological footing. In particular, there is no reason to believe that *S* is an ontologically negative relation that does not refer to any natural kind.[6]

Condition (iii) applied to *S* should not arouse any controversies, as the sphere of the open future cannot include the entire space-time (not all events are open—some have already come to be). As for (iv), its analogue can be presented as the requirement that if *a* is in the causal past of *b*, then *Sab*. Condition (v) is obviously

[5] I owe this objection to the anonymous referee.

[6] Notice that there exists an alternative method of defining the relation of indefiniteness: we can identify it as the converse rather than the complement of the becoming relation (this definition is based on the observation that if *x* is determined as of *y*, *y* had to be undetermined as of *x*). In such a case the transitivity of *S* would follow from the transitivity of *R*.

applicable to the case of the relation S. To complete the entire preparatory procedure, we may wish to finally stipulate that for every two space-time points a and b, either Rab or Sab should hold (but, again, this stipulation is not to be seen as part of the definition of S, but rather as an a posteriori supposition regarding the ontological structure of the world—I defer a discussion of this supposition until a later moment).[7] However, it is now easy to notice that no relations R and S can possibly satisfy all of the above requirements. To see this, let us notice first that for every point b lying in the absolute past of a given point a there is a point c which is at space-like separation relative to both a and c. Since $\sim Rac$ and $\sim Rcb$ (from Stein's theorem), it follows that Sac and Scb. Using the transitivity assumption we derive that Sab, which clearly shows that the openness relation S is universal (since we already know that a stands in the relation S to all points that are either in the absolute future of, or at space-like separation from a), contrary to what we have assumed. Consequently, there can be no relations R and S satisfying all the conditions presented above.

One may view the above rather elementary result as playing directly into the hands of Stein's adversaries. Perhaps they were right that his goal was not attainable after all. However, rather than joining the camp represented by Maxwell, Putnam and Rietdijk, I would like to consider some possible ways of repairing Stein's proposal. It should now be clear that the only way to do this is to abandon some of the requirements imposed on both the definiteness and the openness relations. But which ones? One option that comes to mind is to give up the exhaustiveness condition (characteristically, Stein never mentions this condition *explicitly*). The consequence of such a move would be that for some points a and b neither Rab nor Sab would hold (there is a point b which is neither determinate, nor indeterminate as of a). Now we are in a position to identify relations R and S which would satisfy all the remaining conditions: for a given a Rab would hold iff b is in the causal past of a, and Sac would hold iff c is in the causal future of a. That way we would obtain a division of the entire space-time into three spheres: the absolute past of a (a's past light cone together with its interior), the absolute future of a (the interior and the surface of a's future light cone excluding a itself), and the remaining space-time region containing all points space-like separated from a.

However, the main difficulty with this solution is that the ontic status of points space-like separated from a becomes quite problematic. These points are neither determinate, nor unsettled with respect to a. So what are they? Is there a third state between being open and having passed into being? One may suggest that this is the stage in which things are just in transition: they are literally "passing". However, let us notice that this state of passing from the open future to the determinate past is far from being momentary. If a thing is located sufficiently far from the reference point a, a substantial part of its history (even its whole history)

[7] We have to stress that on Stein's analysis the present belongs to the already definite sphere of reality (every event is definite as of itself). Thus we have a division into two spheres of reality: one that has already come to exist, and the other still unsettled, rather than into the three stages of the temporal growth: the past, the present, and the future. The distinction between these three stages is usually associated with another metaphysical intuition regarding existence in time, according to which only the present exists (is objectively real), whereas both the past and the future are not real (the past *was* real, the future *will be* real). This view, known as "presentism", is not implied by Stein's version of the theory of objective becoming. Rather, Stein's proposal seems to display an affinity with the view which T. Crisp (2003, p. 219) calls "dynamic eternalism", and S. Savitt (2006) dubs "possibilism". One of the examples of this view is C.D. Broad's growing block theory.

can be contained within the spatiotemporal region outside a's future and past light cones. Consequently, some physical objects can remain for quite a while in an ontic "limbo" between being unsettled and being definite. Thus, without some insight into the nature of this third state, we cannot treat the considered solution as a viable one. We must look elsewhere for a weakening of the imposed conditions.

At this point we should turn a critical eye on the requirement of transitivity that was imposed on both the definiteness relation R and the openness relation S. While this condition seems to be deeply rooted in our intuitions regarding time and becoming, not all of those intuitions have to be seen as sacrosanct. As I will argue later, the transitivity requirement put on the definiteness relation R can be seen as a way of securing an even more fundamental intuition regarding the passage of time, which may be called "the cumulativity requirement". However, it turns out that the cumulativity can be preserved without demanding the relation of becoming to be transitive. Thus, the price to pay for rejecting transitivity may be not too high after all.

Stein himself made a remark suggesting that his choice may be to reject the transitivity of the second relation S, while retaining the transitivity of R. Commenting on the passage in which Maxwell argues that the supposition that spacelike separated events are indefinite with respect to each other leads to a contradiction, Stein writes

> Maxwell assumes that the relation "x and y are indefinite for each other" is transitive; but why? I see no compelling grounds for such an assumption; and against it, I see what seems to me a most compelling reason: that it renders the notion of becoming incompatible with the special theory of relativity (Stein, 1991, p. 151).

Indeed, if we allowed that mutual indefiniteness be transitive, we would have the consequence that an event is indefinite with respect to itself, which is not acceptable. But of course the situation in which an event x is indefinite as of y, and at the same time y is indefinite as of x, does not occur in the world of classical physics. Classically, the relation of being indefinite is asymmetric: if x is indefinite as of y, y cannot be indefinite as of x. However, when we consider the world of relativistic physics things change a bit. Here for a given event x we have an entire sphere of events which can neither causally influence x, nor be influenced by it. If we insist that the impossibility of the existence of a (classical) causal link between two events is the mark of indeterminacy, then we should accept that an event spacelike separated from a given one is indefinite as of the latter. However, according to the special theory of relativity, the impossibility of a causal connection goes in both directions. Hence the relation of indeterminacy ceases to be asymmetric (which does not imply that it becomes symmetric—there are still pairs of events for which causality can occur in one direction only). On top of this, indeterminacy is no longer transitive, as it is clearly possible to have three events x, y, z such that x cannot be causally connected with y, y cannot be causally connected with z, but x is causally connected with z. Hence it looks like Stein's analysis can be upheld by accepting the non-transitivity of S.

However, I would like to argue that by allowing the indeterminacy relation S to be non-transitive, we have opened the door to the possibility of applying the same correction to its counterpart—the relation of definiteness R. As I suggested earlier, our pre-relativistic intuitions support the transitivity assumptions equally strongly for both R and S. Once we have made room for violating this intuition in the case of one of these relations, by the same token we should consider the possibility in which the other one ceases to be transitive. And here is why I think this option ought to be given at least some consideration. We can insist that the part of the world history that is ontologically fixed and determined with respect to a given point a should consist of all and only events that *cannot* be causally influenced by what transpires at a. In the classical case this sphere is identical to the region whose inhabitants can, in turn, causally influence a. Thus the two plausible intuitions regarding the fixed past merge together. But this is not the case with relativity. The region that cannot be causally influenced by the state at a consists of the absolute past of a plus the region space-like separated from a. And it should be clear that if we follow this line of thought, we will end up with a relation of definiteness R that is plainly intransitive. As a consolation, we can notice that S—which is the complement of R—becomes in turn transitive. For in this case Sab holds iff b is contained in the future light cone of a.

As a result, we have shown that there are two pairs of relations R_1, S_1 and R_2, S_2 that satisfy Stein's weakened requirements for the theory of objective becoming that is compatible with special relativity. They can be defined as follows:

- R_1ab iff b is located outside the future light cone of a, or $a = b$,
- S_1ab iff b is located inside the future light cone of a, excluding a,
- R_2ab iff b is located inside the past light cone of a,
- S_2ab iff b is located outside the past light cone of a.

Clearly, S_1 is transitive, while R_1 is not, and R_2 is transitive, but S_2 is not. Both conceptions of becoming agree on the ontological status of the spheres of reality contained in a's past and future light cones, but they differ regarding the contentious case of the region space-like separated from it. According to the first approach, these spheres belong entirely to the fixed, definite "past", whereas the second approach deems it part of the open future. As we have argued, each of these solutions can be claimed to preserve some intuitions regarding causal influence that arose in the classical world. Unfortunately, not all of these classical intuitions can be saved.

However, one may still insist that there are additional arguments against accepting the non-transitive relation R_1 as a genuine becoming relation. One of these arguments may be extracted from the 1996 paper by Clifton and Hogarth. They make a point there that we need transitivity "to secure at least a minimum intersubjective agreement between observers on what events have become" (Clifton and Hogarth, 1995, p. 356; the same argument is briefly repeated in Dorato, 2006). Explaining this thought further, they state that "worldline transitivity asserts that any two observers whose worldlines happen to cross at a point must agree, at least at that point, on what events have become real." As Clifton and Hogarth work with a more general notion of becoming with respect to a given worldline,

they indeed have to make sure that two such relations of becoming given with respect to two worldlines will agree at the point of their intersection. According to Clifton and Hogarth the worldline transitivity requirement is the condition stating that if a has become as of b relative to the worldline λ passing through b, and b has become as of c relative to a different worldline λ' passing through c, then a has become as of c relative to λ'. Setting $b = c$, we immediately obtain the required agreement regarding what the two observers (each associated with his respective worldline) passing each other at a given point consider to be their appropriate definite past. But, as Clifton and Hogarth admit, Stein does not consider a worldline dependent becoming, but a universal one. Hence he does not need an independent assumption to ensure the agreement between different observers regarding what events have become real as of a given point. Thus no support for the transitivity requirement can come from this particular angle.

I believe that the transitivity of the becoming relation can be claimed to be required for a different reason. The assumption of transitivity may be seen as securing the fundamental feature of the passage of time which I propose to call "cumulativity". It should be clear that every reasonable theory of objective becoming ought to assume that, for a given observer, the area of the definite past grows as time progresses, and consequently no event considered by this observer to have become as of a given point can later be excluded from the absolute, determinate past. More precisely, the cumulativity requirement would state that for every worldline λ and any points a, b and c such that $b \in \lambda$, $c \in \lambda$, if Rab and Rbc, then Rac. Obviously, this condition follows immediately from the assumption that R is transitive *simpliciter*. Hence, it looks like we have produced an abductive argument in favor of the transitivity of R, as long as we subscribe to the view that cumulativity is an essential feature of R.

I agree that cumulativity had better not be abandoned in a reasonable theory of becoming. However, for cumulativity to be secured we don't need full transitivity! For instance, the relation R_1 defined earlier is not transitive, but is still cumulative. Clearly, if we consider any two points a and b on a given worldline (hence a and b are time-like or null separated) such that a is earlier than b, then the area $A = \{x\colon\ R_1ax\}$, which comprises of the events outside a's future light cone, will be included in $B = \{x\colon\ R_1bx\}$. In other words, the relation R_1 becomes transitive when restricted to a given worldline. The only case in which the transitivity of R_1 is violated is when points a and b are space-like separated, for in such a case we can find a point c space-like separated from a and lying in the absolute future of b. This proves that R_1ac and R_1ba but $\sim R_1bc$. However, in this case there is no observer whose worldline passes through the points a and b. Hence the principle of cumulativity as stated above is not violated.

2.

The special theory of relativity with its principle of Lorentz invariance is not the only modern physical theory that poses a potential threat to the concept of objective becoming. Recently it has become clear that quantum mechanics too can cast

doubts on the intelligibility of the separation of reality into a fixed past and an open future. The reason for this is the existence of entangled quantum states, the most famous example of which is the notorious singlet-spin state. In the case of two spatially distant particles constituting the EPR system there exists a strict correlation between the results of measurements of the same spin-component on both particles, even though the locations of the measurements may be space-like separated. The correlation is such that the outcomes are always opposite: if one of them is "up", the other has to be "down". This fact is often conveniently summarized in the slogan "quantum mechanics violates outcome independence". Dorato (1996) uses the quantum case of outcome-dependence as an argument against Stein's definition of objective becoming that we have just generalized.[8] His argument draws on a remark made earlier by Clifton and Hogarth (1995, p. 385). In a nutshell, their main point is that the existence of the correlation between distant outcomes shows that one outcome acts as a cause (or, at least, as a *partial* cause) of the other outcome. But if there is a causal link between space-like separated outcome-events, it seems that space-like separated events have to be definite as of one another. And this, plus Stein's condition of transitivity of the definiteness relation R, leads to the conclusion that R has to be the universal relation.[9]

Several lines of defense can be considered against such an attack. One option is to question that the correlation between the outcomes revealed in the EPR case is genuinely causal. It has been suggested by some authors, most notably by M. Redhead (1987), that the mere counterfactual dependence between space-like separated outcomes does not warrant the existence of a causal link. According to this approach, causality between two events requires more than only that if one event hadn't happened, the other would not have occurred. Even without plunging into a detailed exegesis of various conceptions of cause we can notice that the case of non-local quantum correlations does not look much like ordinary causation. After all, how often do we encounter a situation in which an event x causes an event y, and y reciprocates, causing in turn x? Such a symmetric causal link not only offends our common sense, but moreover forces us to further choose between rejecting the transitivity of the causal relation and accepting that an event can be its own cause. Neither of these options seems to be particularly attractive, which shows that there is something fishy about "causality" between space-like separated events.[10]

[8] One unfortunate feature of the otherwise very deep and illuminating paper by Dorato is that the author appears to equate the notion of determinateness as used by Stein with the quantum determinateness understood as possessing a sharp value of a given observable. However, it is clear to me that these two notions cannot be equivalent. It is perfectly possible that a quantum system at a given point a does not possess a sharp value of an observable (hence its state is indeterminate in the quantum sense), and yet its state is determinate in Stein's sense as of a different point b, because a is in the causal past of b.

[9] This argument, as it stands, does not seem to directly affect the relation of becoming R_1, according to which space-like separated events *are* mutually definite. However, I take it that the non-local quantum correlations can be as well exploited for the purpose of undermining this conception. For instance, it may be claimed that because a acts as a (partial) cause of a space-like separated event b, b *cannot* be determinate as of a, because whatever has become real as of a does not *need* a cause at a. One may note, though, that the fact that we can put a different spin on the same EPR example in order to produce dramatically different conclusions casts doubts on the legitimacy of this particular quantum counterexample to Stein's conception (see later in the text).

[10] The defenders of the causal character of the EPR connections usually seek support from Lewis's conception of causality based on the notion of counterfactual dependence (Lewis, 1973b). However, it is maybe worth noting that in his latest proposal Lewis substantially modified his counterfactual approach to causality (Lewis, 2000). According to this novel ap-

Unfortunately, the question of whether the correlations between outcomes in the EPR situation deserve to be called "causal" looks suspiciously close to being merely terminological. Some features of the EPR correlations bear unquestionable similarities to ordinary causal relations (e.g. counterfactual dependence), but some differ radically from our standard understanding of causality (for instance, symmetry). Whether we include or exclude quantum non-local correlations from the general kind "causality" depends mostly on what features of causal relations we see as essential. But even if we agree that the EPR outcome-outcome correlations are indeed causal, we may still rebut Clifton, Hogarth, and Dorato's argument for the thesis that two space-like separated outcomes of measurements on a singlet-spin state system are co-determinate. This argument is based on one crucial premise: that if *a* is a cause (either partial, or complete) of *b*, then *a* must have become as of *b*. On the surface, this postulate looks like a consequence of Stein's condition (iv), which ensures that the relation of *causal connectibility* should be included in the relation of becoming. However, Stein had in mind the causal connectibility that obeys the constraints of the special theory of relativity, i.e. that link two events such that one can be connected with the other with the help of a sub-luminal or luminal signal. Now, it is not at all clear that we can extend Stein's condition to the purported causal links between space-like separated events.

Dorato (1996) in his paper presents an extended argument in favor of the claim that the outcome *a* obtained in one wing of the EPR apparatus cannot be seen as "fixing" the other outcome *b* revealed in the distant wing, unless *a* is definite as of *b*. This claim sounds convincing enough, but only as long as we attach the classical meaning to the word "fixing". Once we realize that the outcome *a* in the EPR situation fixes the outcome *b* only in a very Pickwickian sense, since *a* in turn can be equally legitimately seen as being fixed by *b*, the argument begins to crumble.[11] The relation that holds between the two outcomes of the entangled EPR system is very peculiar and difficult to make sense of in classical terms. One can call this relation "causal" if one wishes to, but the mere fact that we use the familiar word "cause" does not legitimize transferring all classical intuitions that we associate with causality to the quantum non-local case. Following this line of thought further, I can even go as far as to let Dorato and the others use the word "determinateness" with respect to the outcomes of measurements in the EPR case. However, in my language I wish to separate the classical notions of causality and determinateness from those of quantum theory. I shall then introduce the terms "*q*-causality" and "*q*-determinateness" in order to refer to the latter. Now, I admit that the outcomes in the EPR case are *q*-determinate as of one another. But I re-

proach, the causal relation is defined as the ancestral of the relation of *influence*, which in turn can be roughly characterized as the relation that holds between two events *C* and *E* iff some counterfactual alterations of the event *C* are followed by alterations of *E*. Now, it should be clear that the EPR case does not satisfy Lewis's 2000 requirement of causality, since no alteration of one outcome (regarding, for instance, its timing or specific manner in which this outcome has been revealed) can produce any difference in the other outcome.

[11] I take it as self-evident that when we normally speak of an event *a* fixing an event *b*, this relation involves two elements playing clearly distinct roles: an object that *fixes* and, hence is already definite, and an object that is *being fixed* and, hence comes to be definite because of the first one. Now, this intuition cannot be applied without contradiction to the quantum case of non-local correlations, as what we have here is an event that both fixes and is fixed by another one. That's why I believe that in this case we can speak of "fixing" only metaphorically (to indicate this, we should use a different term, for instance "*q*-fixing").

main unconvinced that q-determinateness has anything to do with the relation of objective becoming that Stein introduced in his proposal.[12]

However, the EPR case still produces a challenge to the theory of objective becoming, although of a slightly different sort from that raised by Clifton, Hogarth and Dorato. Regardless of the controversial issue of their causal character, the non-local quantum correlations imply the truth of certain counterfactual statements that link space-like separated events. It is an open question whether the existence of such true counterfactuals can be satisfactorily accounted for in a theory that is committed to the objective distinction between an already determinate past and a still open future. In what follows I will try to give a positive albeit tentative answer to that question by producing not one, but two possible ways of accounting for true non-local counterfactuals while accepting a theory of objective becoming. But first I will have to say a few words about the semantics of counterfactual conditionals.

Since the time of the famous analysis of counterfactual conditionals given by David Lewis in (1973a), the following truth condition has been commonly accepted as the "gold standard": for a counterfactual "If it were A, then it would be B" (in short: "$A\Box\rightarrow B$") to be true, B has to be true in the closest possible worlds in which A is true (in short: the closest A-worlds). Although Lewis would most certainly disagree with this move, many authors find it natural to interpret the relation of closeness in such a way that when considering a counterfactual event A occurring at a time t, the closest possible worlds are those that share with the actual world the sphere of already determinate events as of t. In the classical Newtonian case these would be the worlds whose events that occurred earlier than t would be identical to the actual events. However, as we have argued before, in the world of the special theory of relativity there are two ways of interpreting the relation of objective becoming: one using the relation R_1 and the other R_2. Hence, two notions of spatiotemporal counterfactuals, which can be labeled C_1 and C_2, emerge naturally (in the following formulas A describes an event located precisely at a space-time point a):

(C_1) $A\Box\rightarrow B$ is true iff B is true in all possible A-worlds that are identical with the actual world in the entire area outside a's future light cone

(C_2) $A\Box\rightarrow B$ is true iff B is true in all possible A-worlds that are identical with the actual world in the entire area inside a's past light cone

Keeping the part of the world that has already become real seems to be a natural way to consider how things could have been different. All the differences brought about by the counterfactual supposition A should be confined to that part of the world that was yet open as of A. However, it turns out that this can't be done when A refers to one of the two outcomes obtained in the separate wings of the EPR

[12] I can conceive of a different argument in favor of the co-determinateness thesis, which nevertheless seems to me utterly misguided. One may play on the ambiguity of the word "real" in the phrase "to become real" and claim that the events that have not yet become real are simply non-existent. But surely a cause of an existing event cannot itself be non-existent. Hence, even a q-cause must be determinate as of its effect. The fallacy committed here, in my opinion, is a result of failing to notice that the separation between the determinate past and the open future is not identical to the separation between the existing (in the fundamental, Quinean sense) and the non-existing events. Stein's conception, as I see it, allows to freely quantify over both determinate and indeterminate events.

system. Suppose that in the actual world two measurements of the z-component of spin have been performed in space-like separated wings located at points a and b, and the result obtained at a was "+", while the other measurement revealed the value "−". Now, if we consider the possibility that the outcome at a was "−", we have to admit that under this supposition the correlated outcome at b must be changed to "+". But, according to (C_1), the point b, which is space-like separated from a, belongs to the already determinate "past", and as such should be kept intact while evaluating the counterfactual. Thus, (C_1) cannot be applied to the non-local quantum counterfactuals.

The second proposal (C_2) does not suffer a fatal blow from the EPR case, as it permits a possibility that a region space-like separated from a can differ from that of the actual world. However, there is still one minor problem left with (C_2). As it turns out, the possible worlds in which to evaluate the counterfactual $A\Box\rightarrow B$ will contain all and only worlds which share with the actual world the past light cone of a, and in which both outcomes are switched. On the other hand, if we wanted to consider the counterfactual situation in which the outcome at b had been switched, we should choose slightly different worlds: those that share with the actual world the past light cone of b rather than that of a. But shouldn't these two sets of possible worlds—the set necessary for considering the switched outcome at a and the set needed for evaluating the switching at b—be the same? After all, it doesn't really matter which outcome we decide to counterfactually switch—in the end we are left with the two outcomes switched anyway.

Hence, it may be claimed that both truth conditions (C_1) and (C_2) are in need of some corrections in order to be applicable to the case of non-local correlations. Fortunately, this task has been already largely fulfilled. The truth condition (C_1) found its generalization in works of Finkelstein (1999) and Bigaj (2004, 2006). Roughly speaking, the idea is to define an appropriate similarity relation between possible worlds that would reduce to (C_1) in the classical case of singular and local antecedent-events. As the proposal goes, in order to compare two possible worlds with respect to their similarity to the actual world, we have to select in each of them the so-called primary points of divergence, i.e. the space-time points at which some change in comparison to the actual world occurred, and such that no points in their causal past contain any divergence from the actual world. In the next step we should consider the set of points defined as the sum of the interiors of the future light cones of all primary points of divergence. A possible world i will be deemed more similar (closer) to the actual world than a world j if and only if i's set of points thus constructed is properly included in the corresponding set of points of j.

Applying this procedure to the EPR case with two correlated particles we can immediately notice that in the closest A-worlds there will be two primary points of divergence: a and b. Hence, the evaluation of the counterfactual should be done in the worlds which agree with the actual world everywhere outside the two light cones: one originating at a, and the other at b. I suggest that this novel approach can be interpreted as extending the notion of becoming that was embedded in the relation R_1 (and its complement S_1). From the perspective of the space-time point a, not the entire area outside a's future light cone can be included in the already

determinate past as of a. Because the outcome revealed at a is lawfully linked with the event that occurs at the space-like separated location b, whatever counts as yet indeterminate as of b can be arguably seen as indeterminate as of a too. Consequently, due to the non-local link between the physical events transpiring at a and b, the area of the determinate past as of a is the same as the area of the determinate past as of b—namely, the spatiotemporal region outside *both* future light cones of a and b.

The procedure that can generalize the truth condition (C_2) has been described in some detail in Bigaj (2006, Sec. 5.4). Here I will only briefly mention that the idea is to evaluate the counterfactual $A\Box\rightarrow B$ in the possible worlds which extend the closest possible A-worlds selected by the similarity relation introduced above. This extension can be informally characterized as follows: for a given world i we consider all the worlds which share with the actual world the sum of all past light cones of the primary points of divergence in i. Even with this sketchy description of the required procedure it should be clear that in the EPR case the worlds in which to evaluate the counterfactual $A\Box\rightarrow B$ will be those that share with the actual world the interiors of *both* past light cones of a and b. This will immediately ensure that the same worlds should be considered no matter which outcome we decide to counterfactually alter. And, again, it may be claimed that by generalizing the truth condition (C_2) we have extended the notion of objective becoming represented by R_2. Now we may be tempted to admit that the area of events that have become real as of the spatiotemporal point of one measurement should be identical with the sum of *two* past light cones, each with its apex at one of the two measurements.

Stein's proposal regarding objective becoming, as well as its generalization defended in the first part of this paper, is based on one crucial assumption: that whether an event has become as of a given point a depends only on its location relative to a and not on what physically transpires at a. In other words, the division between a determinate past and an open future is a matter of spatiotemporal geometry only—no space-time point should be privileged over other points with respect to its spheres of determinateness and openness. But now we have come to a radically different view of becoming. We have suggested that physical events that transpire at a given point can affect the area of what has already become as of this point. This conception of becoming can be referred to as "dynamical" as opposed to the geometrical approach to becoming defended by Stein. According to the dynamical conception of becoming, the set of events that have become real as of a given point a depends on the lawful correlations that may exist between the physical events happening at a and other events space-like separated from a. Thus, the passage of time may be ultimately dependent on the physical furnishing of space-time (which should hardly come as a surprise, taking into account the well-known lessons from the general theory of relativity).

Returning for a moment to the case of two particles prepared in the entangled singlet-spin state, we can find an additional support for our treatment of becoming in the form of the notion of non-separability. It is often claimed that the primary non-classical feature of quantum entangled systems is not their non-locality but rather non-separability, i.e. the fact that it is impossible to split up the system into

its well-defined components.[13] When we take this into account, it may be easier for us to accept that the objective past of one particle that constitutes the entangled system has to include the past of the other particle as well, since in fact there are no separate particles but one physical system comprising distinct areas of space-time. If fundamentally there is just one object whose non-separable parts occupy separated spatiotemporal regions, then the best we can do in terms of characterizing its area of definite past is to follow Myrvold's suggestion of how to define the sphere of objective becoming for spatiotemporally extended objects. But this procedure will lead us precisely to the sphere of becoming that contains not one, but two past light cones that originate at the locations of the two entangled particles.

In spite of the attractiveness of the dynamical approach to becoming, one has to acknowledge some of its pitfalls as well. One particular problem is related to the requirement of cumulativity that we insisted any theory of becoming should satisfy. When we consider a worldline passing through the point a at which one of the two measurements was performed, we have to admit that from the perspective of some points on this worldline located later than a, the sphere of determinate events will exclude some of the events that were included in it previously as of a. These will be some of the events located outside a's past light cone, but inside b's light cone that were included in a's sphere of becoming as a consequence of quantum non-separability. However, the non-separability vanishes as a result of measurements which disentangle the entangled systems. Hence, after a we are back to the normal sphere of determinateness which is obviously less-inclusive. This is a serious challenge, and at this moment I am unable to find a satisfactory solution.

Another difficulty is created by more sophisticated examples of quantum entangled states, involving for instance not two (as in the EPR case), but three particles. The so-called GHZ state of three particles implies the existence of a strict correlation between the outcomes of measurements of the following form: for three appropriately selected observables with two possible outcomes "+" and "−" each, the number of outcomes "−" must be odd (for details see Greenberger et al. (1990), Maudlin (1994, pp. 24–27)). This implies that when one outcome is considered to be counterfactually switched, precisely one of the other two must be switched too in order to compensate for this change. But the rules of quantum mechanics don't predict which of the remaining two outcomes should be altered. As a consequence, when all three measurements are done at space-like separation, we have to consider two possible worlds in which the outcome for the observable X_1 has been changed: one in which X_2 changed its value but X_3 remained fixed, and the other where the opposite holds. But now the question is: which of these worlds defines the area of the determinate events as of the moment of the first measurement? Clearly, we have two equally convincing answers here. Restricting ourselves to the case of the relation R_2 of objective becoming we can note that there are two possibilities: the area of the events that have become real as of the measurement of X_1 can consist of either the past light cones of the measurements X_1 and X_2, or the past light cones of the measurements X_1 and X_3. But based on the

[13] For an extensive discussion of the notion of quantum non-separability see Howard (1985, 1989, 1997), Healey (2004), Esfeld (2004).

current proposal we are unable to decide which is the case, and this fact seriously undermines the appeal of the dynamical theory of becoming.

My ultimate conclusion is that neither the special theory of relativity, nor quantum mechanics forces us to entirely give up on the idea of objective becoming. In stating this I agree with Stein and the fundamentals of his analysis, although I reject his uniqueness theorem, proposing instead two distinct ways of interpreting becoming that are relativistically invariant. As for the argument from quantum mechanics, I believe that it rests on an unjustified extension of some intuitions that are grounded in the classical notion of causality. Consequently, even if one is unwilling to give some credence to my concept of dynamical becoming, one may still safely apply Stein's analysis of becoming to the case of quantum non-locality. However, it is apt to restate here Dorato's warning that he applied to Stein's original result: the fact that the notion of objective becoming is *compatible* with both relativity and quantum mechanics does not show that *there is becoming in the physical world*. At this moment I neither feel a desire, nor see myself prepared, to take a position on this particular issue.

REFERENCES

Augustynek, Z., 1991. Time. Past, Present, Future. PWN, Kluwer, Warsaw, Dordrecht.

Bigaj, T., 2004. Counterfactuals and spatiotemporal events. Synthese 142 (1), 1–20.

Bigaj, T., 2006. Non-Locality and Possible Worlds. A Counterfactual Perspective on Quantum Entanglement. Ontos Verlag, Frankfurt.

Clifton, R., Hogarth, M., 1995. The definability of objective becoming in Minkowski spacetime. Synthese 103, 355–387.

Crisp, T., 2003. Presentism. In: Loux, M.J., Zimmerman, D.W. (Eds.), The Oxford Handbook of Metaphysics. Oxford University Press, Oxford, pp. 211–245.

Dorato, M., 1996. On becoming, relativity and nonseparability. Philosophy of Science 63, 585–604.

Dorato, M., 2006. Absolute becoming, relational becoming and the arrow of time: Some non-conventional remarks on the relationship between physics and metaphysics. Studies in the History and Philosophy of Modern Physics 37, 559–576.

Esfeld, M., 2004. Quantum entanglement and a metaphysics of relations. Studies in the History and Philosophy of Modern Physics 35B, 601–617.

Faye, J., Scheffler, U., Urchs, M., 1997. Introduction. In: Perspectives on Time. Kluwer, Dordrecht, pp. 1–58.

Finkelstein, J., 1999. Space-time counterfactuals. Synthese 119, 287–298.

Greenberger, D., Horne, M., Shimony, A., Zeilinger, A., 1990. Bell's theorem without inequalities. American Journal of Physics 58, 1131–1143.

Healey, R., 2004. Holism and nonseparability in physics. In: Zalta, E.N. (Ed.), The Stanford Encyclopedia of Philosophy (Winter 2004 Edition), http://plato.stanford.edu/archives/win2004/entries/physics-holism/.

Howard, D., 1985. Einstein on locality and separability. Studies in History and Philosophy of Science 16, 171–201.

Howard, D., 1989. Holism, separability, and the metaphysical implications of the bell experiments. In: Cushing, J., McMullin, E. (Eds.), Philosophical Consequences of Quantum Theory. Reflections on Bell's Theorem. University of Notre Dame Press, Notre Dame, pp. 224–253.

Howard, D., 1997. Space-time and separability: Problems of identity and individuation in fundamental physics. In: Cohen, R.S., Horne, M., Stachel, J. (Eds.), Potentiality, Entanglement and Passion-at-a-Distance. Kluwer, Dordrecht, pp. 113–141.

Lewis, D., 1973a. Counterfactuals. Harvard University Press, Cambridge, MA.

Lewis, D., 1973b. Causation. Journal of Philosophy 70, 570–572.
Lewis, D., 2000. Causation as influence. Journal of Philosophy 97, 182–197.
Maudlin, T., 1994. Quantum Non-Locality and Relativity. Blackwell, Oxford.
Maxwell, N., 1985. Discussion: Are probabilism and special relativity incompatible? Philosophy of Science 55, 640–645.
Myrvold, W.C., 2003. Relativistic quantum becoming. British Journal for the Philosophy of Science 54, 475–500.
Putnam, H., 1967. Time and physical geometry. The Journal of Philosophy 64, 240–247.
Redhead, M., 1987. Non-Locality, Incompleteness and Realism. Oxford University Press, Oxford.
Rietdijk, C.W., 1966. A rigorous proof of determinism derived from the special theory of relativity. Philosophy of Science 43, 598–609.
Savitt, S., 2006. Being and becoming in modern physics. In: Zalta, E.N. (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2006 Edition). http://plato.stanford.edu/archives/fall2006/entries/spacetime-bebecome.
Stein, H., 1991. On relativity theory and the openness of the future. Philosophy of Science 58, 147–167.

CHAPTER 13

Relativity, the Passage of Time and the Cosmic Clock

Peter Forrest*

Abstract

Who's Afraid of Special Relativity? Too many philosophers who shouldn't be. In particular presentists and Growing Block theorists tend to prefer alternatives. The presentist William Craig [Craig, W.L., 2001. Time and the Metaphysics of Relativity. Kluwer Academic Publishers, Dordrecht], for instance, holds a neo-Lorentzian position, and the Growing Block theorist Michael Tooley [Tooley, M., 1997. Time Tense and Causation. Clarendon Press, Oxford] endorses Winnie's theory. I, however, follow Richard Swinburne [Swinburne, R., 1983. Verificationism and theories of space-time. In: Swinburne, R. (Ed.), Space, Time, and Causality. D. Reidel, Dordrecht, pp. 63–78] in holding that whatever objection could once have been based on Special Relativity has been undermined by the discovery of the almost isotropic expansion of the Universe, which provides us with the Cosmic Clock. Neither Presentism nor the Growing Block nor Storrs McCall's falling branches theory [McCall, S., 1994. A Model of the Universe. Oxford University Press, Oxford] are, I say, threatened by Special Relativity, but the Newtonian thesis of a uniform rate of the passage of Time has to be modified.

INTRODUCTION

In this chapter I explicate the familiar idea of the passage of Time and then use this explication to assess the force of the two objections from Special Relativity that I know of: the Objection from the Relativistic Invariance of all laws, and the Epistemic Objection. Let me say at the outset that by an objection from Special Relativity I mean one that can be stated using Special Relativity, not one that assumes Special Relativity is perfectly accurate. So it is in order to consider the relevance of the almost isotropic expansion of the universe, even though that is strictly incompatible with Special Relativity.

* School of Humanities, University of New England, Armidale, Australia

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00013-2

© Elsevier BV
All rights reserved

1. THE PASSAGE OF TIME

For the sake of definiteness, I shall suppose my favourite theory of Time, the Growing Block, according to which the present is the boundary between the real past and the unreal future. Again for the sake of definiteness, I shall suppose that Spacetime is continuous rather than discrete. The passage of Time is most easily illustrated if, nonetheless, the structure of successive presents is discrete. In that case, the history of the Universe consists of successive states, for each of which there is a next state. And in each state the whole of Spacetime includes as a part, usually a proper part, the spacetime corresponding to the previous state. Hence Spacetime grows by adding layers that are usually everywhere of positive thickness. Supposing all this to be so, then the thickness of the layers is the rate at which Time passes, which is therefore measured in seconds, not in seconds per second—which bears on the well known objection to the passage of Time that the rate of passage, a second per second, is trivial.[1] Time passes uniformly, then, if the temporal thickness of successive layers is the same at all places and times.

The rate of passage of Time is, however, ambiguous. For we are considering the rate at which Time understood one way is correlated with Time understood another. It is not clear which way round this is meant to be. Is the rate seconds per layer or layers per second? Suppose the thickness of successive increases tending towards infinity. Then, as a referee pointed out to me, Time would hardly pass at all. In the limiting case the whole of the universe might be in a single layer, which is equivalent to the Block Universe theory. Yet the Block Universe is static and so in that case the rate of passage of time should be zero, not infinity. What this shows, I think, is that the passage of time might as reasonably be measured in layers per second as the other way round. On the other hand, if we thought of some being, God, whose specious present corresponds to a single layer, then the thicker the layers the quicker physical processes would seem to go. Hence there is also a case for measuring the passage of time in seconds per layer. There is no contradiction here, merely a choice of conventions, and I stick to the rate being in seconds per layer.

Unless Spacetime is itself discrete the above account faces a dilemma. Either we take a whole layer of Spacetime as the present or we take only the boundary of reality as the present. In the former case some moments of the present are before others, which is counter-intuitive. In the latter case most of Spacetime is never at any time in the true (as opposed to specious) present, which is also counter-intuitive. The second horn is far sharper given Presentism than it is if we adopt the Growing Block or McCall's Falling Branches theory (McCall, 1994).

I suspect, however, that our intuitions about the present are really about the specious present, not the true present. Provided no layer is thinner than the specious present of any conscious being, there are no moments that were never in the specious present. Considerations of the specious present also blunt the first horn. For within the specious present there is a little bit of B series, with, for example, the successive notes of a melody experienced in order. In so far as our intuitions

[1] The objection is probably due to D.C. Williams, but is only gestured towards in his paper, 'The Myth of Passage' (1951).

about the present are based upon experiencing the specious present there is no intuitive objection to the thesis that the true present is of positive duration, although presumably, less than that of the specious present. In that case the true present, like the specious one, will contain a little bit of B series. The objection comes not from intuition, but I suggest, from a widely accepted theory of Time that posits a present of zero duration. Presumably the reason for that posit is the supposed arbitrariness of any positive duration. But if successive layers are added to reality then their thickness is the duration of the true present, which is not arbitrary.

If readers remain convinced that both horns are sharp, then there is an alternative not subject to this dilemma. It is to assume that the system of successive presents is not discrete but continuous. In that case it constitutes a *foliation* of Spacetime, namely a division of Spacetime into disjoint space-like hypersurfaces. I shall call this the *Absolute Foliation* because any two points on the same hypersurface of the foliation are co-present, that is, they were or are both present together. (I ignore the future as not yet existing.) I call the hypersurfaces of the Absolute Foliation the *hypersurfaces of co-presence*.

We may then say that Time passes *uniformly* with respect to a frame of reference R, if for any hypersurfaces of co-presence S and S', the temporal duration (w.r.t. R) separating any point x on S from any point x' on S', is the same for any x on S and any x' on S'. If the Absolute Foliation consists of hyperplanes, that is flat hypersurfaces, then Time passes uniformly with respect to one frame of reference if and only if it passes uniformly with respect to every other and this occurs if and only if the hyperplanes are parallel.

This hypothesis implies that between any two hypersurfaces of co-presence there are infinitely many others. So it is not possible to assign them integers in such a way that if $n > m$ then the nth is later than the mth. Hence I introduce a real valued parameter, which I measure in tards, such that the later layers are tardier. The rate of passage of time at one place is then measured in seconds per tard. Unfortunately, there are many different choices of parameters that represent the same ordering of the layers. We could, therefore, insist that somewhere, Montreal say, we keep the standard tard, and stipulate what is the rate of passage in seconds per tard in Montreal. To avoid a spurious appearance of triviality I stipulate that in Montreal 2π tards pass per second. Elsewhere the rate is different. Of course it is then a matter of stipulation that time passes at that rate in Montreal, just as it was once a matter of stipulation that the standard metre bar was a metre long if kept at the right temperature. This should put to rest the objection that time passes at a trivial rate. But if more needs to be said I note that light in a vacuum travels at the trivial rate of a second per second too, if we measure distances in (light) seconds. Speed would be trivialised and velocity become merely a matter of direction if everything moved at a second per second everywhere. Likewise there is some force to the triviality objection on the Newtonian thesis that Time passes everywhere at precisely a second per second. But why should it?

2. THE OBJECTION FROM RELATIVISTIC INVARIANCE

It is widely supposed that there is an objection to the passage of Time from Special Relativity. The following is intended as a rigorous version of this objection.

(1) Any given pervasive uniformity in Spacetime almost certainly holds of (nomic or metaphysical) necessity.
(2) The structure of successive presents due to the passage of Time would be a pervasive uniformity in Spacetime.

From (1) and (2):

(3) If Time passes there is a necessary structure of successive presents.

Now

(4) If Special Relativity is correct any necessary structures in Spacetime are relativistically invariant.

But

(5) The structure of successive presents cannot be relativistically invariant.

Hence

(6) Time does not pass.

Premise (1) depends on just what we mean by 'pervasive'. In 'Kneale's argument Revisited', George Molnar (1954) considers the truth that there is no river of Coca Cola. And he supposes—plausibly enough—that this is true not just on Earth now but at all times and places. Presumably such generalisations do not count as pervasive uniformities. I characterise a pervasive uniformity as the analysis that neo-Humeans give of natural necessity, that is, a generalisation is a pervasive uniformity if it would be treated as necessary given a neo-Humean account of necessity such as David Lewis' development of Ramsey's theory (Lewis, 1973: 72–77). Hence the 'almost certainly' qualification in Premise (1) is intended only for anti-Humeans, such as myself, who take pervasive uniformity as extremely good evidence for necessity, but not as entailing it. While such details are of independent interest I do not think the application of Premise (1) to a uniform system of successive presents, as in the Newtonian conception of Time, is problematic.

Premise (4) follows if we state Special Relativity as the conjunction of: (4a) It is a law of nature that the electromagnetic constant c has a fixed value in cm per secs, with (4b) All laws of nature are invariant with respect to changes from one frame of reference to another moving relative to the first with some uniform velocity less than c. Stating it that way is to get metaphysical, since scientific theories as such deal in pervasive uniformities not laws of nature, but because of Premise (1) this is not problematic.

Premise (5) might seem vulnerable to Howard Stein's (1968) suggestion that the present could be shaped like the surface of a light cone. But my response is that although Lorentz invariance is the interesting part of relativistic invariance there is also the rather boring translation invariance. Light cones have vertices and so

are not translation invariant, even though a light cone centered on the origin is Lorentz invariant. Likewise hyperbolic hypersurfaces, with coordinates given by the equation $\sqrt{(t^2 - x^2 - y^2 - z^2)} = k$ for varying positive k, although Lorentz invariant are not translation invariant. A translation-invariant family of successive presents must be a system of parallel hyperplanes and these are not Lorentz invariant. So Premise (5) holds.

My reply to the Objection from Relativistic Invariance is that Premise (2) is incorrect. The actual system of successive presents is not the pervasive uniformity. It is the fact that Time passes uniformly that is the uniformity, and hence necessary. But this does not contradict Special Relativity. All it implies is that, whatever the system of successive presents is, any relativistic transformation of this system would also be a (nomologically) possible system of successive presents. Within the scope of the assumption of Special Relativity we may suppose that the successive presents are parallel hyperplanes. Then it would be a law that Time passes in such a way that *some* system of parallel hyperplanes are successive presents, but not a law as to which one is. Once a given system of parallel hyperplanes has established itself, then the law tells us that all subsequent presents are also parallel, but how it got established is something to do with initial conditions, or, more accurately something to do with the early stages of the universe when Special Relativity was not a good approximation.

To be sure the Newtonian theory of Space and Time does contradict Special Relativity. For it states that Spacetime is just Space, on the one hand, and Time, on the other, and so has a *necessary* product structure as point/moment pairs. Hence there is a necessary system of successive presents, namely those parameterised by the Time coordinate. And this system fails to be invariant in Special Relativity.

My reply to the objection, then, is that Special Relativity does not render even the uniform passage of Time problematic, and any who think it does have confused the uniform passage of Time with the point/moment pair structure of Spacetime. I shall argue below that Time does not pass uniformly but does so to a good approximation. This will, therefore, be superficially inconsistent with my reply to the Objection from Relativistic Invariance. I am not abashed. My reply was stated, like the objection, within the scope of the assumption of Special Relativity. For the remainder of the chapter I merely assume that Special Relativity is a good approximation.

There is, however, a rejoinder to my reply: I owe the objector an explanation of the origin of just one system of successive hyperplanes—or approximations thereto—as the system of successive presents. Well, I say, as debtors do, just give me a while and I shall repay. But first I consider the other objection.

3. THE EPISTEMIC OBJECTION

This concerns what I call the *inscrutability of the Absolute Foliation*. It is said to be problematic to posit an account in which there is a fact as to which pairs of distant past events are co-present, without us having any way of discovering it. But just

what is problematic? John Mackie talked of the 'Conspiracy of Silence' (1983: 21). A nice phrase, but it hardly helps us say what the problem is. Adrian Heathcote has suggested that this structure of successive hypersurfaces would be the sort of thing that physics aims to discover but resistant to any scientific attempt to detect it, directly or indirectly.[2] Hence physics would not merely be actually incomplete, but incompletable. This would be like a theory with a fundamental constant, call it kappa, that must not be positive but is otherwise undetectable. Annoying, I grant, but why epistemically objectionable?

Here is my attempt at stating the Epistemic Objection. Time seems to pass, so there is a debate between projection theorists and realists about the passage of Time. The former consider the belief that Time passes and the associated belief that the future is unlike the past to be projections onto reality of a subjective sense of the passage of Time. The latter, of whom I am one, start from Reid's Principle of Credulity, the presumption in favour of realism about the way things seem. Projection theorists argue that the presumption is defeated because they have a simpler theory. And those philosophers who believe in the passage of Time but nonetheless grant that Special Relativity raises a problem may be interpreted as agreeing that the projection-theorist has the simpler theory but as denying that this advantage is enough to defeat the initial presumption in favour of realism. I therefore interpret the existence of an inscrutable Absolute Foliation as problematic because—and only because—it is the mark of greater complexity.

The nicest way of meeting the Epistemic Objection would be if quantum theory in fact does enable us to discover co-present events as a result of entanglement.[3] Quantum theory is, however, notoriously hard to interpret, so I shall ignore it. Instead I refine the reply given by Richard Swinburne (1983: 73) that the expansion of the universe provides us with a clock that gives us an accurate enough guide to co-presence, because if we use a frame of reference with respect to which the expansion is almost isotropic then we may specify which past events are co-present. Call this the Cosmic Clock Defence.

Before developing the Cosmic Clock Defence I note two relevant points about the expanding universe. The first is that it is not a basic law of nature that any universe must expand in an almost isotropic fashion. Rather the almost isotropic expansion of our universe is due to the laws of nature *together* with conditions in the early universe. For if it had been a basic law of nature then we would expect the law to require perfectly isotropic expansion, which would prevent the formation of galaxies. There is, therefore, the project of explaining why the conditions of the early universe resulted in this almost isotropic expansion. Currently there are three hypotheses: divine providence; the anthropic selection effect together with a plurality of universes, and the inflationary Big Bang. The first two are based on the requirement that the expansion be fairly isotropic but not perfectly so if the universe is to be suited to life. The chief difficulty with them is that the deviations from perfect isotropy are many orders of magnitude less than the minimum

[2] In the discussion following a paper I gave on the Growing Block theory at Sydney University in 2005. This is also one of the points made by Simon Saunders in his unpublished 2000 paper, 'How Relativity Contradicts Presentism'.

[3] Suppose two particles have opposite spin states but without a determinate direction of spin. Then we might suppose observing, and so making determinate, the spin of one particle simultaneously results in a determinate spin for the other.

required for a life-friendly universe. So we have here an embarrassing case of overtuning. (See Forrest, 2007: 98–99.) Tentatively, therefore, I endorse the inflationary Big Bang hypothesis, according to which there was a period during which the universe expanded at ever increasing rates (Mukhanov, 2005, ch. 5).

This first point, namely the case for the Inflationary Big Bang, will be brought to bear on the second point, which is that there are two distinct hypotheses that should be clearly distinguished: Relativistic Isotropic Expansion and Absolute Isotropic Expansion. The former states that there is some (approximately special relativistic) frame of reference X with respect to which the expansion is (almost) isotropic. The latter states in addition that, for any such frame of reference X, the hypersurfaces of co-presence of the Absolute Foliation are good approximations to the constant time hyperplanes specified by X. If a case can be made for the Absolute Isotropic Expansion then the Cosmic Clock Defence is vindicated. But the weaker Relativistic Isotropic Expansion hypothesis does not directly meet the Epistemic Objection. For if the hypersurfaces of co-presence of the Absolute Foliation are not good approximations to the constant time hyperplanes specified by X then we do not know which events are co-present.

It is tempting to make an inference from Relativistic Isotropic Expansion to Absolute Isotropic Expansion on the grounds that it is simpler to hypothesise an isotropic expansion rather than one that is more rapid in some direction than in the opposite (cf. Swinburne, 1983: 73). But that appeal to simplicity would be justified only if it were a basic law of nature that any universe expands in an isotropic fashion, in which case we should believe the simpler, although stronger, law as well as the weaker one. As it is we should go back to the inflationary Big Bang hypothesis to decide whether it supports the stronger Absolute Isotropic Expansion or just the weaker Relativistic Isotropic Expansion. Maybe there is some quantum theoretic argument that the almost isotropic expansion is due to the simultaneous decoherence of a previously coherent state and maybe quantum theory requires that there be a privileged foliation of Spacetime into Spaces. I have already noted that as the ideal reply to the Epistemic Objection. But it might instead turn out that quantum theory requires no privileged foliation. In that case the details of the inflationary Big Bang would become irrelevant. For then there could be no laws that would introduce the privileged frame of reverence, so it is only the weaker Relativistic Isotropic Expansion hypothesis that would be supported. In any case, it is prudent to assume the weaker hypothesis given the speculative character of the details of the Inflationary Big Bang.

There is more work to be done, then. I need to find a plausible hypothesis that enables us to derive Absolute Isotropic Expansion from Relativistic Isotropic Expansion. In that way I shall have replied to the Epistemic Objection. In addition I shall have paid the debt I previously acknowledged, namely explaining how one system of successive hypersurfaces comes to be that of the successive presents. I shall use the inflationary Big Bang for this purpose, in spite of its tentative character, on the grounds that it is the most plausible explanation for the almost isotropic expansion. But first I need to expound the passage of Time in a relativistically invariant way.

4. MAKING THE PASSAGE OF TIME RELATIVISTICALLY INVARIANT

The hypervolume of a region of Spacetime is a relativistic invariant.[4] Now, on the Growing Block theory the universe grows in the temporal dimension as well as expanding spatially. Accordingly I propose the Law of Uniform Growth stating that the 'growing block' grows uniformly not in duration so much as in hypervolume. More precisely, if we consider two hypersurfaces of co-presence in the Absolute Foliation then any light cone with a vertex on one of them extending as far as the other encloses the same hypervolume as any other such cone. This is the relativistic analog of the Newtonian uniform passage of Time. If the expansion of the universe were perfectly isotropic then this uniform growth would imply the strict Newtonian thesis of uniform passage of Time. But given that the expansion is merely almost isotropic it follows that Time passes almost but not perfectly uniformly.

Suppose that at some time in the past the Absolute Foliation corresponded to a frame of reference R' in which the universe did not expand in an isotropic way, so to an observer at rest relative to R' the background radiation is significantly blue-shifted in one direction and red-shifted in the opposite direction. And consider an observer E at rest relative to another frame R with respect to which there was almost isotropic expansion. Then there were two points far distant from the observer that although co-present would, according *to the observer E*, be in the distant past and the distant future, respectively. Therefore, if the rate of expansion of the universe was increasing, the light cones that E considers in the distant past are thinner and those that E considers in the distant future are fatter. The proposed Law of Uniform Growth states that nonetheless these cones have the same hypervolume. Therefore the hypersurfaces of the Absolute Foliation will have to be separated by a greater temporal duration in those regions that E takes to be in the distant past and less in those regions that E takes to be the distant future. The effect of this is to tilt the hypersurfaces of co-presence more and more until eventually they come to approximate the constant t hypersurfaces of the frame R, with respect to which there is almost isotropic expansion. Notice, however, that if the rate of expansion of the universe were decreasing the reverse effect occurs and so any deviation from isotropy would increase.

The most plausible hypothesis to explain Relativistic Isotropic Expansion is the inflationary Big Bang, which posits a rapidly increasing rate of universe expansion. Combining this with the Law of Uniform Growth we obtain the Absolute Isotropic Expansion hypothesis. For even if in the early stages the Absolute Foliation deviated from one for which isotropic expansion held the inflationary expansion would reduce this deviation. Hence Absolute Isotropy also holds, vindicating the Cosmic Clock defence.

[4] Quibble: to ensure invariance under improper Lorentz transformations we need to consider the signed hypervolume.

CONCLUSION

The Argument from Relativistic Invariance may be used against the thesis that Spacetime can be factored into Space and Time, but not against the passage of Time understood as the growth of Reality, or even against the Newtonian conception of Time as passing uniformly of necessity. The Epistemic Objection requires this Newtonian conception be revised so that it is not Time that passes uniformly but the growing block that grows uniformly. The inflationary Big Bang then justifies not merely the Relativistic but also the Absolute Isotropy hypothesis, completing the Cosmic Clock Defence.

This defence of the passage of Time may be adapted to other dynamic theories of Time, if we say that the hypervolume of History increases uniformly. I find it more plausible, however, that a fundamental necessary truth be stated as the uniform growth of Reality rather than the uniform growth of History.

ACKNOWLEDGEMENT

This chapter was originally titled 'Who's Afraid of Special Relativity?' I would like to thank all who contributed to the discussion, and especially to the referee. I would also like to thank the Australian Research Council for a Discovery Grant that funded the research.

REFERENCES

Forrest, P., 2007. Developmental Theism: From Pure will to Unbounded Love. Oxford University Press, Oxford.

Lewis, D., 1973. Counterfactuals. Blackwell, Oxford.

Mackie, J.L., 1983. Three steps towards absolutism. In: Swinburne, R. (Ed.), Space, Time, and Causality. D. Reidel, Dordrecht, pp. 3–22.

McCall, S., 1994. A Model of the Universe. Oxford University Press, Oxford.

Molnar, G., 1954. Kneale's argument revisited. Phil. Rev. 78, 79–89.

Mukhanov, V., 2005. Physical Foundations of Cosmology. Cambridge University Press, Cambridge.

Stein, H., 1968. On Einstein–Minkowski space-time. Journal of Philosophy 65, 5–23.

Swinburne, R., 1983. Verificationism and theories of space-time. In: Swinburne, R. (Ed.), Space, Time, and Causality. D. Reidel, Dordrecht, pp. 63–78.

Williams, D.C., 1951. The myth of passage. Journal of Philosophy 48, 457–472.

CHAPTER 14

Time and Relation in Relativity and Quantum Gravity: From Time to Processes

Alexis de Saint-Ours*

Abstract

We examine the question of spatialized time in physics and the tension between timelessness and essential temporality. We recall that these issues were analysed by process philosophers, in particular Bergson and Whitehead, and we show the recurrence of this debate in quantum gravity through Christian's Heraclitean generalization of relativity versus Barbour's Platonia. We then argue that a relational account of time, such as Rovelli's, does not picture a changeless world but change without time, that is a world of processes in which dynamics does not refer to an external and fictitious parameter t but is intrinsically built into the systems. We end by a presentation of Rovelli's relational Quantum Mechanics stressing the fact that both Quantum Physics and General Relativity lead to a very general relational framework. We recall that some philosophers of the French tradition of epistemology have laid down building blocks for understanding this framework. We conclude by proposing a metaphysics of dynamical relationalism for a future theory of quantum gravity.

> "The history of natural philosophy is characterized by the interplay of two rivals philosophies of time—one aiming at its "elimination" and the other based on the belief that it is fundamental and irreducible." G.J. Whitrow (1980).

1. ESSENTIAL TEMPORALITY AND TIMELESSNESS IN PHYSICS

1.1 Introduction

Is time rooted in the very nature of reality or a mere stubborn illusion? As recalled by Griffin (1986), time as experienced involves three characteristics: (1) a cate-

* University of Paris VIII and Laboratory "Pensée des Sciences", École Normale Supérieure, Paris, France

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00014-4
© Elsevier BV
All rights reserved

gorical distinction between past, present and future, (2) constant becoming and (3) a one-way irreversible direction. The past is fully determined, the present is inherently becoming and the future is undetermined. Regarding these essential properties, closed past, creative present and open future,[1] time in classical physics, mainly in Newtonian physics and in classical electrodynamics, appears totally irrelevant to time as we experience it. Indeed, it is well known that, with the exception of the second law of thermodynamics, the equations of classical physics do not give evidence to time's irreversibility, quite the reverse. This gap between physical time and experienced time has led to very different positions. For some thinkers, the timelessness of classical physics should make us realise that experienced time is an illusion as for others it shows physics' incompleteness in taking into account the world's essential temporality. In this dilemma, one will recognize no more than another manifestation of the old controversy between Heraclitus and Parmenides. In the 20th century, with the birth of relativity theory, Quantum Mechanics (QM) and chaos theory, the actors have changed but the dilemma between essential temporality and timelessness has remained more vivid than ever.

1.2 Spatialized time and process philosophy

The critique of the timelessness of physics was formulated in different perspectives at the end of the 19th and at the beginning of the 20th century by two process philosophers: Whitehead and Bergson. This initial critique found deep echoes later in the century in the work of de Broglie, Milič Čapek, Olivier Costa de Beauregard and Abner Shimony.

Bergson's philosophy is an attempt to show that physics profoundly misunderstands the nature of time and never deals with authentic time: what Bergson calls duration. With duration the French philosopher means the essence of time: he thinks that time's main attribute is invention, that is "*Continuous creation of unpredictable novelty*". Commentators have named this principal characteristic of duration *virtuality*. In the second chapter of *Time and Free Will: An essay on the Immediate Data of Consciousness*, in order to show the difference between space and duration, Bergson sets up a major distinction between quantitative multiplicities and qualitative multiplicities. On one hand, you have space and number. Space is homogeneous, quantitative and actual. On the other hand, you have duration which is in total opposition to space. Duration is heterogeneous, qualitative and virtual. In order to understand duration as invention, Bergson resorts to a distinction between two couples: actuality and virtuality on one side; possibility and reality on the other. Virtualities become actual and possibilities are realised. There is a relation of resemblance between possibility and reality whereas the actual does not resemble the virtuality it is incarnating. The latter explains why creation is duration's essential attribute and the idea of creative present and open future.

Bergson's main thesis is that when physics talks about time, it does not talk about duration but about a very poor conception of time, what he calls spatialized time. Spatialized time is the measurable time, symbolized by the variable t that

[1] "This coming into existence involves the transformation of potentialities into actualities" (Griffin, 1986, p. 3).

occurs in physical formulae. It can not shed any light on the true essence of time as it is the ghost of space. Spatialized time is a quantitative multiplicity in which duration has been eliminated. The variable t has no relation to real time at all. This is striking if one considers the representation of time as an horizontal line: this idea of closed past, creative present and open future has disappeared from this representation. Duration can not be symbolized by a line since a line is actual and duration virtual. Let us recall here that much of the work of Prigogine was to reveal the existence of irreversibility in physics at the macroscopic and microscopic level, while stressing his agreement with Whitehead and Bergson and "*seeing his own task to be that of giving scientific content and precision to their metaphysical speculations*" (Griffin, 1986, p. 17).

Bergson's critique was constructed against the timelessness of classical physics and formulated few years before Einstein's special theory of relativity. In 1922, Bergson publishes *Duration and Simultaneity*, a book in which he compares his own conception of time to time in special relativity. In this book, Bergson wrongly tries to show that contrary to Einstein's interpretation, there is absolute simultaneity and absolute time. The French philosopher criticized relativity and Minkowskian spacetime for having invented a new way of spatializing time. In spite of this false understanding of special relativity,[2] this critique of spatialized time in relativity (and in physics in general) might be of crucial importance for current attempts in quantum gravity. Indeed, in his last book, *The Trouble with Physics*, Lee Smolin analyses the different theories aiming at the unification of General Relativity (GR) and QM and writes:

> "I believe there is something basic we are all missing, some wrong assumption we are all making. [...]. My guess is that it involves two things: the foundations of quantum mechanics and the nature of time. [...]. More and more, I have the feeling that quantum theory and general relativity are both deeply wrong about the nature of time. It is not enough to combine them. There is a deeper problem, perhaps going back to the origin of physics.
> Around the beginning of the seventeenth century, Descartes and Galileo both made a most wonderful discovery: You could draw a graph, with one axis being space and the other being time. A motion through space then becomes a curve on the graph. In this way, time is represented as if it were another dimension of space. Motion is frozen, and a whole history of constant motion and change is presented to us as something static and unchanging. If I had to guess (and guessing is what I do for a living), this is the scene of the crime.
> We have to find a way to *unfreeze time*—to represent time without turning it into space. I have no idea how to do this. I can't conceive of a mathematics that doesn't represent a world as if it were frozen in eternity. It's terribly hard to represent time, and that's why there's a good chance that this representation is the missing piece" (Smolin, 2006b, pp. 256–257).

[2] It must be said in order to defend Bergson, that at that time even professional physicists made interpretational mistakes in their understanding of Einstein's theory.

Can we represent time without turning it into space? This question is closely related to other problems: is time relational or substantival in nature? What is the connection between time and change? We believe that one of the keys on this road to get rid of spatialized time in physics is to understand time as built into the systems—as opposed to time as an external (and sometimes fictitious) parameter. Relational time fits this conception since it gives a picture of change without referring to a variable t. This might sound paradoxical but we will try to make clear that relational time gets rid of fictitious time and fictitious dynamics and shows the way to understand time as process.[3]

Before the examination of this question, let us say a few words about the status of time in special relativity.

1.3 Being and becoming in special relativity

The question of spatialized time in special relativity was notably addressed by two thinkers influenced, although in different ways, by Bergson: Olivier Costa de Beauregard and Milič Čapek. Both focus on the question of being and becoming in relativity. In many articles,[4] Olivier Costa de Beauregard has argued that special relativity pictures a timeless *and* changeless world.[5] The argument is well known and was defended by various philosophers and physicists. Indeed, it was long ago stated by Weyl when he claimed that "*the objective world simply* is, *it does not* happen" (Weyl, 1949, p. 116). Costa de Beauregard (a student of de Broglie) argues that in the context of Minkowski's four dimensional spacetime, everything is already written and that change is relative to human's perception as a lack of not being able to perceive four dimensions. Roughly, the idea is that the motion of a point in time is represented by a stationary curve in a four dimensional spacetime. Becoming in three dimensions is an illusion of *being* in four dimensions. Time is spatialized and the world is static, changeless and timeless.[6] In this block universe framework, the concept of lapse of time determines the possibility of change. We have to mention that Olivier Costa de Beauregard agrees partially with Bergson, since he believes, as Bergson argues, that the spacetime of special relativity is a spatialization of time but he differs radically from the French philosopher in trying to show that this is not a lack of Minkowski's representation but a true understanding and representation of the world.

Against this interpretation two arguments can be opposed. First, entropy as a guarantee of time's irreversibility remains unchanged in the context of special relativity. Second, a difference between time and space appears in Minkowski's metric.[7] Relying on such arguments, Milič Čapek has argued that "*the relativistic union of space with time is far more appropriately characterized as a* dynamization of space *rather than a spatialization of time*" (Čapek, 1966) and that the potentiality of the future is preserved in this framework, in accordance with Bergson's understanding of time as the creation of unpredictable novelty. Milič Čapek's idea is that

[3] Further on in this work, we will define processes as changes without time, i.e. changes relative to other changes.
[4] See for example, Costa de Beauregard (1966).
[5] We emphasize this point since as we will show, it is possible to have change without time.
[6] The absence of change is not a necessary consequence of spatialized time.
[7] $ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2)$.

process is an essential feature of reality and is far from being an illusion of a more fundamental changeless world. These issues are still very controversial. We don't want to pursue the status of being and becoming in special relativity. We refer the reader to the analyses of Richard T.W. Arthur (2006), Mario Bunge (1968), Dennis Dieks (2006a, 2006b), Vesselin Petkov (2005 and especially 2007) and Steven Savitt (2006). Let us just mention that this idea of dynamization of space (rather than spatialization of time) was put forward by Milič Čapek because he believed that the last word of Einstein's theory was not Minkowski's spacetime but rather GR which truly realises (as we will see, especially in the framework of Hamiltonian GR) this idea of dynamization of space.

2. THE RECURRENCE OF THIS DEBATE IN QUANTUM GRAVITY

2.1 Time in Quantum Mechanics, Quantum Field Theory and General Relativity

Time plays a problematic role in the framework of the canonical approaches to quantum gravity.[8] This should not come as a surprise once one accepts the initial incompatibility between time in QM and Quantum Field Theory (QFT) on one hand, and time in GR on the other.

Indeed, time in QM is substantival in nature. The parameter t that is involved in Schrödinger's equation,

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi,$$

is an absolute external parameter. It is not described by an operator: this would be in contradiction with the boundedness of the energy. As noted by Alfredo Macías and Hernando Quevedo: "*It is not a fundamental element of the scheme, but it must introduced from outside as an absolute parameter which coincides with the Newtonian time. Since there is no operator which could be associated with time, it is* not *an observable*" (Macías and Quevedo, 2007, p. 44). Consequently, time is not a dynamical variable in QM.

In QFT, the picture is similar, since absolute time is replaced by a set of times (the times of special relativity) that refer to Minkowski spacetime. Since this spacetime constitutes a background, it is substantival in nature and the time variable is

[8] Among the different attempts to unify GR and QM, one can distinguish:

- The canonical approaches that make use of the Hamiltonian formalism, in which spacetime is foliated and an appropriate canonical variable is chosen. This program is a direct quantization of Hamiltonian GR. In the spirit of GR, it does not assume a background spacetime. According to the choice of the variable, one can specify different subclasses of quantization:
  - The oldest version is quantum geometrodynamics in which the canonical variable is the three dimensional metric.
  - Since the end of the 1980s, this approach is pursued in a different form based on ideas of Abhay Ashtekar in which the canonical variable is a three connection. This has led to Loop Quantum Gravity (LQG).
- Covariant approaches. This line of research has led to string theory and M-theory.
- Sum-over-histories line of research.
- Others, such as: twistor theory, non-commutative geometry, causal sets, ...
- For more details on the different attempts, see Macías and Quevedo (2007), Kiefer (2007) and Rovelli (2004, Appendix B) for a historical account.

a background parameter: "*Instead of the absolute Newtonian time, we now have a different parameter associated to each member of the distinguished class of inertial frames. The two absolute concepts of Newtonian physics, i.e. space and time, are now replaced by the single concept of spacetime. Nevertheless, in special relativity spacetime retains much of the Newtonian scheme. Although it is not possible to find an absolute difference between space and time, spacetime is still an element of the quantum theory which does not interact with the field under consideration. That is to say, spacetime remains as a background entity on which one describes the classical (relativistic) and quantum behaviour of the field*" (Macías and Quevedo, 2007, p. 45). Spacetime in QFT is an external non dynamical entity like the absolute time of QM.

In GR, there is a drastic difference with QM and QFT, since here time is dynamical, local and does not constitute a fixed background: "*In GR, space-time is dynamical and therefore there is no absolute time. Space-time influences material clocks in order to allow them to show proper time. The clocks, in turn, react on the metric and change the geometry. In this sense, the metric itself* is *a clock*" (Kiefer, 2007, p. 137). Time in GR is not an external parameter.[9] GR does not possess a naturally preferred time variable whereas QM and QFT do possess such a preferred time.

2.2 The Wheeler–De Witt equation

As explained by Wald (1984),[10] there are two reasons for developing a Hamiltonian formulation of GR. The first reason is that it expresses the dynamical nature of Einstein's equation:[11] "*the viewpoint that Einstein's equation describes the evolution of the spatial metric, h_{ab}, with "time" is perhaps best motivated via the Hamiltonian formulation*" (Wald, 1984, p. 450). The second reason is the desire to obtain a theory of quantum gravity.

The canonical quantization method, when applied to Hamiltonian GR, leads to the Wheeler–De Witt equation, a second order functional differential equation, sometimes considered as the wave function of the universe. As mentioned earlier, the canonical quantization method can be done in two different but closely related perspectives. In quantum geometrodynamics, spacetime (M, g) is sliced into spatial hypersurfaces Σ_t associated with a preferred time. The three dimensional metric on each Σ_t is then used as the appropriate dynamical variable for the canonical formalism. Let Ψ be the wavefunction of the universe. A quantum state of the universe is then a normalizable complex functional on the configuration space. The Wheeler–De Witt equation states that $H\Psi = 0$. One sees that although this equation is a dynamical equation, no parameter t appears on the right side (as it is the case, for example, in Schrödinger's equation). The Wheeler–De Witt equation does not depend on an external time. It is the main dynamical equation of the theory but it makes no reference to time, even more: "*all the quantities entering it are defined on the 3-dimensional hyper-surface Σ_t. This is one of the most obvious manifestations of the problem of time in general relativity. The situation could not be worse! We*

[9] We will come back to the question of time in GR in Section 2.4.4.
[10] Appendix E "Lagrangian and Hamiltonian Formulations of General Relativity".
[11] See also Chapter 21, "Variational principle and initial-value data", of Misner et al. (1970).

have a quantum theory in which the main dynamical equation can be solved without considering the evolution in time" (Macías and Quevedo, 2007, p. 45). It is worth noting that the Wheeler–De Witt equation is ill defined in quantum geometrodynamics whereas a well defined version of the equation has been constructed in the context of LQG.[12] The disappearance of time in the equation has been interpreted in very different directions.[13] Some researchers have claimed the necessity to reintroduce time into the quantum theory by means of auxiliary physical entities conceived as an internal time, whereas others have concluded that quantum gravity leads us to a timeless picture of nature, even more drastic than the situation of time in classical physics. Like others, we think that it is not time that has disappeared from the framework of canonical quantum gravity but the parameter t. We believe that this situation inaugurates a shift from fictitious time in physical theories to the discovery and understanding of change without time, i.e. process. Before turning to this and to the way this understanding of time is encapsulated in a relational theory of time, we would like to mention the work of some thinkers and physicists that have claimed that the timeless situation of canonical quantum gravity is unacceptable regarding time's properties as described in Section 1.1.

2.3 Christian's Heraclitean generalization of relativity

One of them is Joy Christian, a student of Abner Shimony, who recently proposed an original modification of special relativity in order to capture the flow of time and the Heraclitean idea of becoming as a major and non illusory characteristic of reality. He believes that a future theory of quantum gravity must encapsulate the idea of becoming: "*If, however, temporal becoming is indeed a genuinely ontological attribute of the world, then no approach to quantum gravity can afford to ignore it. After all, by quantum gravity one usually means a* complete *theory of nature. How can a complete theory of nature be oblivious to one of the most immediate and ubiquitous features of the world? Worse still: if temporal becoming is a genuine feature of the world, then how can any approach to quantum gravity possibly hope to succeed while remaining in total denial of its reality*?" (Christian, 2007, p. 10).

In this theory, according to our everyday experience and unlike in special relativity's block universe interpretation, the future segments of the worldlines do not "pre-exist" for all eternity and "*we perceive the events in our lives to be occurring* non-fatalistically, *one after another, causing our worldline to "grow", like a tendril on a wall*" (Christian, 2007, p. 6).

Joy Christian pictures this idea in the following way (Christian, 2007, p. 7), see Figure 14.1.[14]

This Heraclitean generalization of special relativity is constructed by introducing the inverse of the Planck time at the conjunction of special relativity and Hamiltonian mechanics. It is based on two postulates:

(i) A modified relativity principle in a pseudo-Euclidean space ε made up of Minkowski spacetime M and an internal space of states N.

[12] See Rovelli (2004).
[13] See Butterfield and Isham (2006), Kuchař (1992, 1999) and Rickles (2006).
[14] I thank Marina for the drawing of the figure.

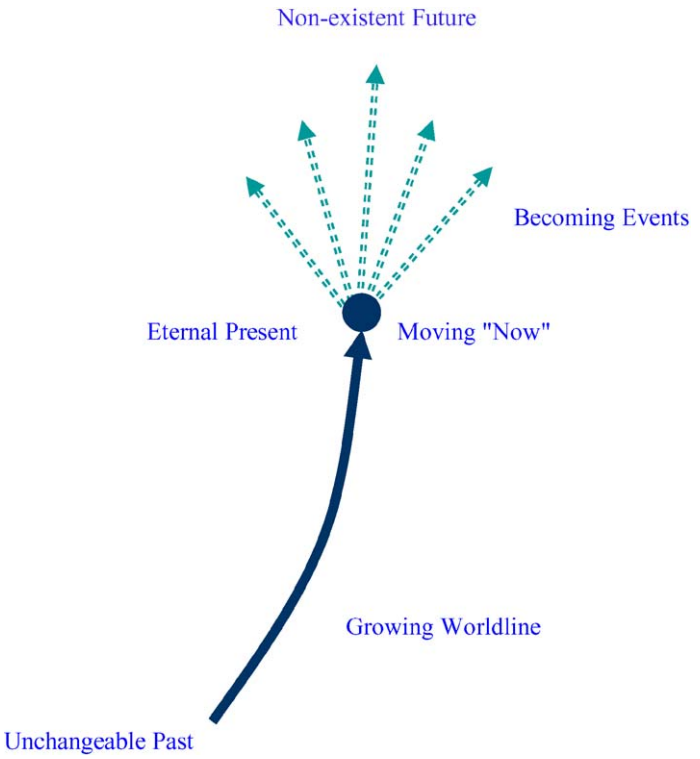

FIGURE 14.1

(ii) No time rate of change of a dimensionless physical quantity can exceed the inverse of the Planck time.

This leads to a new invariant (of the ε space): $d\tau^2 = dt^2 - c^{-2}dx^2 - t_p^2 dy^2$.

In this theory, place is a function of time and state, time a function of place and state and state a function of time and place:

$$\begin{cases} x = x(t,y) \\ t = t(x,y) \\ y = y(t,x) \end{cases}$$

The state dependence of time produces the necessity of becoming in the theory. The theory naturally captures the flow of time and its essential becoming. As one can see, the theory tries to grasp time's essential attributes (closed past, moving and creative present, open future), as pictured in the above diagram.

2.4 Relational physics

2.4.1 Relational theories of time and space

In a completely opposite direction, Julian Barbour has developed a timeless theory of physics which is also an implementation in physics of Mach's relational ideas

about time and space. Before turning to Julian Barbour's theory, let us say a few words about relational theories of space and time.

In "Physical Time: The Objective and Relational Theory",[15] Mario Bunge distinguishes four theories of time:

- Kant's in which time is absolute and subjective.
- Newton's in which time is absolute and objective.
- Berkeley's in which time is relational and subjective.
- And finally, Lucretius' theory in which time is relational and objective.[16]

Here, a relational theory of time is a theory in which time does not exist by itself. It is anchored in something else, usually but not always change, which then becomes more basic than time. An objective theory of time is a framework in which time is a primary feature of the world and does not pertain to the cognitive subject. Mario Bunge explains that a relational theory of time can or cannot be in accordance with Einstein's special relativistic framework. In other words, one has to distinguish between a relational theory of time and a relativistic theory of time. Among the precursors of relational time, Mario Bunge sees: Aristotle (time does not exist by itself since it is the measure of motion), Lucretius,[17] Leibniz (time is the order of successions) and Mach. As we will see, a radical relational and objective theory of time claims its total disappearance. In this view, time is just a convenient artifice or a redundant concept. This is the position of Julian Barbour and, with some differences, Carlo Rovelli. It was also advocated by Mach: "*According to Mach, in any statement containing the variable t, the latter can be replaced by a reference to some phenomenon dependent on the earth's angle of rotation*" (Bunge, 1968, p. 370).

Analogously, a radical relational and objective theory of space claims its disappearance: "*A reformulation is suggested in which quantities normally requiring continuous coordinates for their description are eliminated from primary consideration. In particular, space and time have therefore to be eliminated, and what might be called a form of Mach's principle must be invoked: a relationship of an object to some background space should not be considered—only relationships of objects to each other can have significance*" (Penrose, 1969, p. 151).

2.4.2 Barbour's Platonia

Julian Barbour argues that the Wheeler–De Witt equation pictures a timeless *and* changeless world: "*I consider a strategy for the reconciliation of GR with quantum theory (QT). This is based on an analysis of the essential structure of the two theories and a consideration of what remains of this structure if, as argued in [The timelessness of quantum gravity I], time is truly non-existent in the kinematic foundations of both theories. I suggest that quantum gravity is* static *and simply gives relative probabilities for all the different possible three-dimensional configurations the universe could have*" (Barbour,

[15] I thank Vincent Bontems for the reference of this article.

[16] In this paper, inspired by Russell and Reichenbach (see Reichenbach, 1999, II.3 "The Causal Theory of Time"), Mario Bunge proposes an axiomatic theory of objective and relational time.

[17] "Time itself does not exist. It gets meaning from things, from the fact that events are in the past, or that they are now or that they will happen in the future. It must not be claimed that anyone can sense time by itself apart from the motion of things or their restful immobility" De Natura Rerum.

1994b, pp. 2875–2876). Much of Julian Barbour's work has been done to understand how motion and time can emerge from a static world (Platonia)[18] by means of his theory of "time capsules", fixed patterns that encode history and motion. Very naturally, this theory is related to the implementation of a Machian program not only in relativity but also in a reformulation of classical physics.[19]

Barbour has achieved this in a joint work with Bruno Bertotti. They have elaborated a relational theory of Newtonian dynamics, with no reference to absolute space or time, in which only relative distances occur. In this *intrinsic dynamics*: "*there exists a uniquely distinguished parameter that does make the equations simple and has identical properties to Newton's absolute time. However, it is not introduced independently but as a very natural weighted average of all the motions in the universe. This is a very satisfactory result. For it suggests that the mysterious invisible time that seems to control all motions in our part of the universe is simply determined by the average of all the motions of the universe*" (Barbour, 2006, p. 93). Similarly, Julian Barbour has showed how classical GR can be understood in such a timeless way.

Julian Barbour did more than anyone else to revive this relational approach to physics. Doing so, he had a great influence on Lee Smolin and Carlo Rovelli. We believe that their approach, slightly different from Barbour's, is a new way of understanding time, not as the external independent variable t but as a true process. We believe that this shift from time to process is a decisive step in the attempt to get rid of spatialized time.

2.4.3 Rovelli's and Smolin's relationalism

It is striking to notice that the founders of LQG are driven by deep conceptual motivations. In his recent work, Carlo Rovelli insists that the problem of quantum gravity will not be solved unless physicists and philosophers reconsider questions such as: What is space, what is time, what is the meaning of position or the meaning of motion? The heart of these conceptual issues lies in the long standing debate between relationalism versus substantivalism. Is space or spacetime an entity, a stage or a convenient name for the relationship between physical entities?

If space is to be considered as an entity, then one has to accept that space exists by itself and that physical entities move *in* space. Both Newton's absolute space and Minkowski spacetime are substantival in this way. A logical consequence of this conception is that motion and position are relative *to* space. In Rovelli's and Smolin's relational perspective, however, space is just a convenient name for labeling relationships between physical entities. Position and motion of a physical entity are not to be referred to an absolute stage but have to be considered relative to other physical entities. As Rovelli (2004) explains, space is no more than the "touch", the "contiguity" or the "adjacency" relation between objects. Physicists often talk about "background independence" in this context. As Lee Smolin explains: "*The debate between philosophers that used to be phrased in terms of absolute versus relational theories of space and time is continued in a debate between physicists who argue about background dependent versus background independent theories*" (Smolin,

[18] See Barbour (1994a, 1994b) and Barbour (2004).
[19] See Barbour (2004).

2006a, p. 204). Smolin argues that the relational strategy is more explanatory and more easily falsifiable, as becomes clear, e.g., in the Leibniz/Clark debate.

GR is relational because spacetime location is relational: "*The point is that the only physically meaningful definition of location within GR is relational. GR describes the world as a set of interacting fields including* $g_{\mu\nu}(x)$*, and possibly other objects, and motion can be defined only by positions and displacements of these dynamical objects relative to each other*" (Rovelli, 2001, p. 108). But it is partly relational since dimension, topology, differential structure and signature are fixed and constitute some kind of background.[20]

The way of taking into account GR's conceptual issues has deep consequences for the different attempts of finding a quantum theory of gravity. For instance, in perturbative string theory, a background spacetime is reintroduced. For the loop theorists but also for many others,[21] this is not acceptable because they consider that GR's main lesson is the disappearance of spacetime as an entity. These physicists stress that the future theory of quantum gravity will have to take this into account. This is why in their view quantum gravity has to be background independent. It is evident to loop theorists that we cannot detach technical aspects of GR from their conceptual or foundational aspects: GR is this diffeomorphism invariant theory in which spacetime is relational.

Since in GR spacetime and the gravitational field are the same entity, Carlo Rovelli argues that there is no spacetime but just the gravitational field. This reinforces his relational philosophy in which physical entities are particles and fields. There is no background and the properties of the elementary particles and fields consist entirely in the relationships among them. These relationships evolve, and time is nothing but the parameter ordering the change in the relationships.[22] In this relational view, one has to accept that we don't live *in* space and that we don't evolve *in* time either.

A first glance at these ideas might lead someone to think that this situation describes once again an unacceptable timeless world. But we argue that what has disappeared from this theory is the *parameter* t. We have seen that the parameter t occurring in most of physics' equations is a spatialized non-dynamical background parameter. It simulates time in the equations but does not capture its essence. Since change remains in Rovelli's and Smolin's conceptual framework, one can see it as a true understanding of time or better, of process. We believe that the relational shift inaugurated by Julian Barbour and pursued by Rovelli and Smolin leads to a timeless physics which truly grasps processes. The world is not made of things evolving in time, it is made of processes:[23] "*But relativity and quantum theory each tell us that this is not how the world is. They tell us—no, better, they scream at us—that our world is a history of processes. Motion and change are primary. Nothing* is, *except in a very approximate and temporary sense*" (Smolin, 2001, p. 53).

[20] See Smolin (2006a, p. 205).

[21] "Thus, in the covariant perturbation approach to formulating a quantum theory of gravity, it appears that meaningful physical predictions cannot be made. In addition, this approach has a number of other unappealing features. The breakup of the metric into a background metric which is treated classically and a dynamically field γ_{ab}, which is quantized, is unnatural from the viewpoint of classical general relativity" (Wald, 1984, p. 384).

[22] "The relationships are not fixed, but evolve according to law. Time is nothing but changes in the relationships, and consists of nothing but their ordering" (Smolin, 2006a, p. 204).

[23] In HGR7 (2005), in Las Canarias, I remember Bill Unruh writing on the board: When is a particle?

2.4.4 Process: Change without time

Carlo Rovelli has proposed a theory for this relational understanding of time. As we have seen, in a substantival account of time, time is an exterior and unobservable parameter, as for example time in the context of Newtonian physics. In this case, dynamics always appears to be inauthentic and time spatialized. In other words, in a substantival framework, motion and dynamics are not really dynamical because time is an external and unobservable entity. However, Rovelli's relational account of time does not picture a changeless world. On the contrary, it pictures a world of authentic processes and events in which dynamics does not refer to an external and fictitious parameter t but is intrinsically built into the systems. In this framework, the world is inherently dynamical: "*Thus, a general relativistic theory does not deal with values of dynamical quantities at given spacetime points: it deals with values of dynamical quantities at "where"-s and "when"-s determined by other dynamical quantities*" (Rovelli, 2007b, p. 1310).[24]

As stressed by Julian Barbour and Carlo Rovelli, the disappearance of the time coordinate in the Wheeler–De Witt equation has nothing to do with quantum gravity *per se* since it has already disappeared in the Hamilton–Jacobi formalism of GR. Retrospectively, this should make us realise that the disappearance of time in the equation of canonical quantum gravity is not a novelty since a timeless Newtonian dynamics and a timeless GR can also be constructed.[25]

This relational account of time is what Carlo Rovelli calls "physics without time". He argues that in the context of GR and of quantum gravity, time as an external parameter is a superfluous hypothesis. In *Quantum Gravity* (Sections 1.3.1; 2.3.2; 2.4.4; 3.1; and 3.2.4), he points out that GR predicts correlations between physical variables but is not about physical variables with respect to a preferred time t.[26] This is also the case in classical mechanics. In other words, there is change without time. Carlo Rovelli recalls the following story about Galileo. The Italian founder of classical physics was in the Pisa cathedral watching the oscillations of the grand chandelier. Galileo had the intuition that this motion was isochronous. To check this, he used his pulse as a clock and noticed the isochronism of the chandelier. A few years later, doctors used pendulums to measure human pulses. One might first say that this is paradoxical. But not at all! What this anecdote shows is that evolution is not about the change of variables with respect to time, but about changes with respect to other dynamical variables. Evolution in classical mechanics deals with dynamical variables with respect to other dynamical variables. When one compares this set of dynamical variables, one can easily check that these observations fit with evolution in t: "*In particular, it gives us confidence that to assume the existence of the* unobservable *physical quantity t is a useful and reasonable thing to do. Simply: the usefulness of this assumption is lost in quantum gravity. The theory allows us to calculate the relations between observable quantities, such as A*(*B*),

[24] As explained by Carlo Rovelli, this is also the case in pre-general relativistic context. But Newton's theory (as special relativity) makes a distinction between dynamical variables and the clocks and rods that measure the background space(time). In GR, this distinction is lost.

[25] On the role of symplectic reduction in this framework, see Belot (2007), Butterfield (2007) and Souriau (1969).

[26] Rovelli (2007b), argues that in the context of GR, one must distinguish between the dynamics of matter interacting with a given gravitational field that determines a local notion of time and the dynamics of the gravitational field itself. In the latter, Carlo Rovelli reminds that there is no external time variable that plays the role of an evolution variable.

$B(C)$, $A(T_1)$, $T_1(A)$, . . . , which is what we see. But it does not give us the evolution of these observable quantities in terms of an observable t, as Newton's theory and special relativity do. In a sense, this simply means that there are no good clocks at the Planck scale. [. . .] The theory[27] *is conceptually well defined without making use of the notion of time. It provides probabilistic predictions for correlations between the physical quantities that we can observe. [. . .]. Thus, there is no background "spacetime", forming the stage on which things move. There is no "time" along which everything flows. The world in which we happen to live can be understood without using the notion of time*." (Rovelli, 2004, pp. 30–31).

Rovelli (2007a) makes a distinction between:

- partial observables: a physical quantity with which one can associate a measuring procedure leading to a number;
- complete observables: a quantity whose value can be predicted by the theory (this definition refers to classical theory but has a quantum equivalent in which the probability distribution of the quantity can be predicted by the theory).

Relying on this distinction, Carlo Rovelli argues that at the fundamental level, the variable t is on the same footing as any other partial observables. Connes and Rovelli (1994) have proposed an interesting theory that describes how a macroscopic notion of time (with its properties of closed past, moving present and open future) emerges from this "timeless" picture.

3. PROCESS AND RELATION

3.1 Rovelli's relational quantum mechanics

The concept of relation also appears in Carlo Rovelli's papers about the interpretation of QM. There is a tradition[28] of relational interpretations of QM: e.g., such interpretations already occur in the work of David Finkelstein, Mermin and Mioara Mugur-Schächter.

In his 1997 paper, Carlo Rovelli argues that in the context of QM, and by analogy with Einstein's rejection of absolute simultaneity as the clue to the physical understanding of the Lorentz transformation, one should reject the notion of absolute (or observer independent) state of a system. By abandoning such a notion, one gets a weaker notion of state in which the state becomes relative to other physical systems: "*Quantum mechanics can therefore be viewed as a theory about the states of systems and values of physical quantities relative to other systems. [. . .]. I suggest that in quantum mechanics, "state" as well as "value of a variable"—or "outcome of a measurement" are relational notions in the same sense in which velocity is relational in classical mechanics. [. . .]. In quantum mechanics* all *physical variables are relational*" (Rovelli, 1997, p. 6).

Carlo Rovelli's idea is that the concept of an observer-independent state of a system is inappropriate at the quantum level. The rejection of absolute state of a system relies on the idea that in QM, different observers may give different

[27] LQG.

[28] See Jammer (1974) and Bitbol (forthcoming) for an analysis of C. Rovelli's relational on.

accounts of the same sequence of events. As Einstein did in 1905 with his derivation of the Lorentz transformation based on the principle of relativity and on the constancy of the speed of light, Carlo Rovelli suggests and tries to implement an analogous manoeuvre in QM. Starting with this idea that all physical variables are relational, he tries to show that QM can be derived from a set of three postulates and two hypotheses.

HYP 1: All physical systems are equivalent, nothing distinguishes macroscopic systems from microscopic ones.
HYP 2: QM is complete.

He underlines that this way of understanding QM relies on information theory as it was developed by Shannon. In this context, information is nothing else than the measure of the number of states of a physical system.

Carlo Rovelli starts by showing that the classical distinction between observer and observed system should disappear. Analysing a sequence of events from two different points of view, the one of the observer and the one of a system external to the measurement, he concludes that two different observers give different accounts of the same sequence of events. As already said, this leads him to the conclusion that the notion of state is not absolute but rather observer dependent. In this context, the notion of an absolute state of a system is replaced by the relational notion of information that a physical system may possess on a system.

This strategy is really the one of a relativist, analogous to how Einstein used the universality of the principle of relativity to give an account of the Lorentz transformation and to construct special relativity. In his analogous attempt, Carlo Rovelli explains that: "*Rather than backtracking in front of this observation*,[29] *and giving up the commitment to the belief that all systems are equivalent, I have decided to take this experimental fact at its face value, and consider it as a starting point for understanding the world*. If different observers give different descriptions of the state of the same system, this means that the notion of state is observer dependent. *I have taken this deduction seriously, and have considered a conceptual scheme in which the notion of absolute-observer independent state of a system is replaced by the notion of information about a system that a physical system may possess*" (Rovelli, 1997, p. 15).

To finish this brief presentation of relational QM, let us mention the three postulates used in the derivation:

POS 1: There is a maximum amount of relevant information that can be extracted from a system
POS 2: It is always possible to acquire new information about a system
POS 3: Superposition principle[30]

There are two levels in Carlo Rovelli's relational QM. One of those levels is a shift whereas the other one is a translation. The first level consists in shifting every absolute sentence into a relational one. It is a relativistic shift. For example, instead of saying that the electron has spin up, one should say that the electron has

[29] The observation that two different observers give different accounts of the same physical set of events.
[30] The key clue is to understand that there is no way to compare the information possessed by *O* with the information possessed by *P* without considering a quantum physical interaction or a quantum measurement between the two.

spin up relative to the Stern Gerlach apparatus. The second level consists in translating every relational sentence into the language of information theory. Why? Because information theory gives form, formalizes the relational philosophy. Information theory appears in this context as the scientific form of the relational intuition.

In a recent paper,[31] Carlo Rovelli and Matteo Smerlak develop a relational interpretation of EPR in which they try to show that within the context of relational QM, it is not necessary to abandon locality. They argue that from the relational perspective, quantum non-locality is an illusion that arises from disregarding the quantum nature of all physical systems.

3.2 Spacetime relationalism and quantum relationalism

Unlike others, such as Penrose, Rovelli claims that there is no interrelation between the interpretation of QM and the quantum theory of gravity. Nevertheless, it is tempting to speculate, as Carlo Rovelli does, about this concept of relation that one finds both in GR and in QM. In *Quantum Gravity*, Carlo Rovelli writes: "*I close with a very speculative suggestion. As discussed in Section 2.3, the main idea underlying GR is the relational interpretation of localization: objects are not located in spacetime. They are located with respect to one another. In this section, I have observed that the lesson of QM is that quantum events and states of systems are relational: they make sense only with respect to another system. Thus, both GR and QM are characterized by a form of relationalism. Is there a connection between these two forms of relationalism?*" (Rovelli, 2004, p. 157). He proposes that there might be a connection between on one hand GR's relationalism, depending on contiguity, and on the other QM relationalism, depending on interaction.[32] There is a connection between contiguity and interaction since systems can interact only if they are contiguous. This is locality. Carlo Rovelli therefore suggests that locality ties together GR's relationalism and QM relationalism and that it might be interesting to develop the idea that contiguity derives from the existence of quantum interaction.

3.3 Relationalism in the French tradition of epistemology

We agree that quantum gravity leads to a metaphysics of relations. It is interesting to note that one finds unknown or forgotten elements of such a metaphysics in the French tradition of epistemology, in particular in the philosophy of Gaston Bachelard (1884–1962) and Gilbert Simondon (1924–1989). The relational aspect of knowledge appears in the structuralist philosophy that emerges in the 20th century in linguistics with Jakobson, and in anthropology with Levy-Strauss. The structuralist philosophy states that the meaning of an element emerges in its relation to another element. Compare what Rovelli says in his 1997 paper: "*The relational aspect of knowledge is one of the themes around which large part of western*

[31] Rovelli and Smerlak (2006).

[32] Since the properties of a given system are relative to another system with which it is interacting, QM relationalism depends on interaction

philosophy has developed. [. . .]. I find the fact that quantum mechanics, which has directly contributed to inspire many of these views, has then remained unconnected to these conceptual developments, quite curious" (Rovelli, 1997, p. 19).

Indeed, Gaston Bachelard and Gilbert Simondon have developed their highly original relational philosophies as a result of reflection about QM and GR. In 1931, in *Noumène et Microphysique*, Bachelard draws relational metaphysical consequences of QM. He argues that quantum theory leads to the idea that there are no substantial properties.[33] In a book on relativity published in 1929, *La valeur inductive de la relativité*, Bachelard stresses the relational aspect of Einstein's theory.[34] A chapter of Bachelard's book is a comment on Born's *Einstein's theory of relativity*. In this book, Max Born describes the process of *relativisation*. As Maurice Solovine, French translator of many of Einstein's books explains, the French language does not have an equivalent of the German *relativieren*, which means to put in relation.[35] With respect to the concept of gravity, Bachelard shows its initial substantival content and its progressive evolution to a purely relational concept. What Bachelard says is not only that there is a philosophical impact of relativity. He says something much more universal. Indeed he states that we have to become *"relativistically" minded*.[36]

Following Bachelard but also Bergson, Simondon has built up a very original philosophy in which he claims there is a need to differentiate *being* from *being an individual*. Simondon argues that being as being is more basic and precedes being as an individual. In other words, individuality is not given in advance, it is the result of a process. This being as a being is therefore said to be *preindividual*. Individuals result from individuation. Simondon's philosophy is an attempt to catch the process of individuation at three different levels:

- at the level of physics
- at the level of life
- and at the level of society

This philosophy is in total opposition to *substantialism* in which being is given at all times and therefore is not the result of a process. But it is also in opposition to *hylomorphism*, that claims, like Aristotle, that individuals are the encounter of a matter (*hylê*) and a form (*morphê*). In *hylomorphism*, form exists before matter, which means that there is no process since form and matter exist before their encounter. In analogy with thermodynamics, Simondon thinks that preindividual beings are like metastable systems. A metastable system has a fragile equilibrium.

33 "La substance de l'infiniment petit est contemporaine de la relation" (Bachelard, 1931, p. 13). "Il convient de retenir que le plan nouménal du microcosme est un plan essentiellement complexe. Rien de plus dangereux que d'y postuler la simplicité, l'indépendance des êtres, ou même leur unité. Il faut y inscrire de prime abord la Relation. Au commencement est la Relation." (Bachelard, 1931, p. 18).

34 For an analysis of Bachelard's book, see Alunni (1999).

35 "Nous avons été obligé de forger ce mot (Relativation), dit M. Solovine (L'Ether et la théorie de la Relativité) pour traduire le mot allemand Relativierung, qui exprime admirablement bien la pensée d'Einstein mais qui n'a pas d'équivalent dans la langue française." (Bachelard, 1929, p. 100).

36 "Dans les prolégomènes de la Relativité apparaît au contraire le besoin de se référer à l'externe, de solidariser en quelque partie la qualité d'un objet avec la qualité de l'objet de comparaison, bref d'expliquer par la référence même." (Bachelard, 1929, p. 103).

Therefore individuation is a process because primitive being is metastable. Simondon's idea is that the distinction between subject and object does not exist before their relation. In other terms, they are produced by a relation which is individuation. This is a major aspect of this philosophy: there is relationalism because there is individuation. And process (individuation) cannot be understood without relationalism (called *realism of relation* by Simondon[37]). Simondon's system is an original account in between Bergson's philosophy of process (that is duration) and Bachelard's relational understanding of QM and GR.

Simondon's philosophy of individuation posits a realism of relation. In *L'individu et sa genèse physico-biologique*, that was published in 1964, the French thinker develops a relational account of QM that relies partly on de Broglie's double solution. In this book, Simondon proposes the following. He considers that there are two ways of understanding the wave particle duality. On one hand, one can, as Bohr does, consider physical individuals as spatially limited entities. Simondon perceives a flaw in the Copenhagen interpretation since it starts with the classical way of representing individuals to finally come to the conclusion that individuals are: "*unsharply defined individuals within space-time limits*". On the other hand, such contradictions disappear if one starts with relational account of individuals.

3.4 Dynamical relationalism

Simondon proposed a highly original program that ties together inherently dynamical systems and relationalism. He stressed the importance of considering relations as giving rise to individuation. We believe that quantum gravity leads to a framework of relationalism and processes,[38] that is a metaphysics of dynamical relationalism in which preindividuated physical entities are individuated by means of physical interaction. We have found hints of this metaphysics in Rovelli's and Smolin's work. In a very "Simondonian" style, John Stachel, in "Structure, Individuality, Quantum Gravity", argues that Quantum Gravity will be dynamically relational because spacetime points in GR and particles in quantum theory fall into a peculiar kind of relationalism in which the relation dynamically individuates un-individuated entities: "*Whatever the ultimate nature(s) (quiddity) of the fundamental entities of a quantum gravity theory turn out to be, it is hard to believe that they will possess an inherent individuality (haecceity) already absent at the levels of both general relativity and quantum theory. So I am led to assume that, whatever the nature(s) of the fundamental entities of quantum gravity, they will lack inherent haecceity, and that such individuality as they manifest will be the result of the structure of dynamical relations in which they are enmeshed*" (Stachel, 2006, p. 58).

[37] "les véritables propriétés d'un être sont au niveau de sa genèse, et, pour cette raison même, au niveau de sa relation avec les autres êtres", Simondon (1964). "Individuation et relation sont inséparables; la capacité de relation fait partie de l'être, et entre dans sa définition et dans la détermination de ses limites: il n'y a pas de limite entre l'individu et son activité de relation ; la relation est contemporaine de l'être; elle fait partie de l'être énergétiquement et spatialement", Simondon (1964).

[38] Changes without time.

ACKNOWLEDGEMENTS

I would like to thank Juan Ferret with whom much of this work was elaborated and discussed, Charles Alunni for his support and precious advices, Carlo Rovelli for his stimulating encouragements and the referee for his helpful comments and suggestions.

REFERENCES

Alunni, C., 1999. Relativités et puissances spectrales chez Gaston Bachelard. In: Revue de synthèse, Pensée des sciences. Albin Michel.

Arthur, R.T.W., 2006. Minkowski spacetime and the dimensions of the present. In: Dieks, D. (Ed.), The Ontology of Spacetime. Elsevier, Amsterdam.

Bachelard, G., 1929. La valeur inductive de la relativité. Vrin.

Bachelard, G., 1931. Noumène et microphysique. In: Etudes. Vrin (1970).

Barbour, J., 1994a. The timelessness of quantum gravity: I. The evidence from the classical theory. Class. Quantum Grav. 11, 2853–2873.

Barbour, J., 1994b. The timelessness of quantum gravity: II. The appearance of dynamics in static configurations. Class. Quantum Grav. 11, 2875–2897.

Barbour, J., 2004. The End of Time. Phoenix (1999).

Barbour, J., 2006. The development of Machian themes in the twentieth century. In: Butterfield, J. (Ed.), The Arguments of Time. Oxford University Press.

Belot, G., 2007. The representation of time and change in mechanics. In: Butterfield, J., Earman, J. (Eds.), Handbook of the Philosophy of Science. Philosophy of Physics. Elsevier–North-Holland, Amsterdam.

Bitbol, M., forthcoming. Physical relations or functional relations. To appear as a chapter of: M. Bitbol, De l'intérieur du monde.

Bunge, M., 1968. Physical time: the objective and relational theory. Philosophy of Science 35 (4), 355–388.

Butterfield, J., 2007. On symplectic reduction in classical mechanics. In: Butterfield, J., Earman, J. (Eds.), Handbook of the Philosophy of Science. Philosophy of Physics. Elsevier–North-Holland, Amsterdam.

Butterfield, J., Isham, C., 2006. On the emergence of time in quantum gravity. In: Butterfield, J. (Ed.), The Arguments of Time. Oxford University Press.

Čapek, M., 1966. Time in relativity theory: Arguments for a philosophy of becoming. In: Frazer, J.T. (Ed.), The Voices of Time. Braziller, New York.

Christian, J., 2007. Absolute being vs relative becoming. arXiv: gr-qc/0610049 v1;
In: Vesselin, P. (Ed.), Relativity and the Dimensionality of the World. Springer, 2007.

Connes, A., Rovelli, C., 1994. Von Neumann algebra automorphisms and time versus thermodynamics relation in general covariant quantum theories. Class. and Quantum Grav. 11, 2899.

Costa de Beauregard, O., 1966. Time in relativity theory: Arguments for a philosophy of being. In: Frazer, J.T. (Ed.), The Voices of Time. Braziller, New York.

Dieks, D. (Ed.), 2006a. The Ontology of Spacetime. Elsevier.

Dieks, D., 2006b. Becoming, relativity and locality. In: Dieks, D. (Ed.), The Ontology of Spacetime. Elsevier, pp. 157–176.

Griffin, D.R. (Ed.), 1986. Physics and the Ultimate Significance of Time. State University of New York Press.

Jammer, M., 1974. The Philosophy of Quantum Mechanics. John Wiley.

Kiefer, C., 2007. Quantum Gravity. Oxford University Press.

Kuchař, K., 1992. Time and interpretation of quantum gravity. In: Kunstatter, G., Vincent, D., Williams, J. (Eds.), Proc. 4th Canadian Conf. on General Relativity and Relativistic Astrophysics. World Scientific.

Kuchař, K., 1999. The problem of time in quantum geometrodynamics. In: Butterfield, J. (Ed.), The Arguments of Time. Oxford University Press, Oxford.
Macías, A., Quevedo, H., 2007. Time paradox in quantum gravity. In: Fauser, B., Tolksdorf, J., Zeidler, E. (Eds.), Quantum Gravity. Mathematical Models and Experimental Bounds. Birkhäuser.
Misner, C.W., Thorne, K.S., Wheeler, J.A., 1970. Gravitation. W. H. Freeman and Company.
Penrose, R., 1969. Angular momentum: an approach to combinatorial space-time. In: Bastin, E.A. (Ed.), Quantum Theory and Beyond. Cambridge University Press.
Petkov, V., 2005. Relativity and the Nature of Spacetime. Springer.
Petkov, V. (Ed.), 2007. Relativity and the Dimensionality of the World. Springer.
Reichenbach, H., 1999. The Direction of Time. Dover (1956).
Rickles, D., 2006. Time and structure in canonical gravity. In: Rickles, D., French, S., Saatsi, J. (Eds.), The Structural Foundations of Quantum Gravity. Clarendon Press.
Rovelli, C., 1997. Relational quantum mechanics. arXiv: quant-ph/9609002 v2.
Rovelli, C., 2001. Quantum spacetime: what do we know? In: Callender, C., Huggett, N. (Eds.), Physics Meets Philosophy at the Planck Scale. Cambridge University Press.
Rovelli, C., 2004. Quantum Gravity. Cambridge University Press.
Rovelli, C., 2007a. Partial observables. arXiv: gr-qc/0110035 v3.
Rovelli, C., 2007b. Quantum gravity. In: Butterfield, J., Earman, J. (Eds.), Handbook of the Philosophy of Science. Philosophy of Physics. Elsevier–North-Holland, Amsterdam.
Rovelli, C., Smerlak, M., 2006. Relational EPR. arXiv: quant-ph/0604064 v1.
Savitt, S., 2006. Being and becoming in modern physics. Stanford Encyclopaedia of Philosophy.
Simondon, G., 1964. L'individu et sa genèse physico-biologique. Presses Universitaires de France.
Smolin, L., 2001. Three Roads to Quantum Gravity. Basic Books.
Smolin, L., 2006a. The case for background independence. In: Rickles, D., French, S., Saatsi, J.T. (Eds.), The Structural Foundations of Quantum Gravity. Clarendon, Oxford.
Smolin, L., 2006b. The Trouble with Physics. Houghton Mifflin Company.
Souriau, J.-M., 1969. Structure des Systèmes Dynamiques. Dunod.
Stachel, J., 2006. Structure, individuality and quantum gravity. In: Rickles, D., French, S., Saatsi, J. (Eds.), The Structural Foundations of Quantum Gravity. Clarendon Press.
Wald, R.M., 1984. General Relativity. The University of Chicago Press.
Weyl, H., 1949. Philosophy of Mathematics and Natural Science. Princeton University Press.
Whitrow, G.J., 1980. The Natural Philosophy of Time. Clarendon.

FURTHER READING

Bergson, H., 2001. Time and Free Will: An essay on the Immediate Data of Consciousness. Dover. First edition: Essai sur les données immédiates de la conscience, 1888.
Butterfield, J. (Ed.), 1999. The Arguments of Time. Oxford University Press.
Butterfield, J., Earman, J. (Eds.), 2007. Handbook of the Philosophy of Science. Philosophy of Physics. Elsevier–North-Holland, Amsterdam.
Callender, C., Huggett, N., 2001. Physics Meets Philosophy at the Planck Scale. Cambridge University Press.
Čapek, M., 1961. The Philosophical Impact of Contemporary Physics. D. Van Nostrand Company.
Christian, J., 2006. Passage of time in a Planck scale rooted local inertial structure. arXiv: gr-qc/0308038 v4.
Frazer, J.T. (Ed.), 1966. The Voices of Time. Braziller.
Lucretius, 1995. De Natura Rerum, J. Kamy-Turpin, éd. bilingue. Aubia, Paris.
Price, H., 1996. Time's Arrow and Archimedes' Point: New Directions for the Physics of Time. Oxford University Press.
Rickles, D., French, S., Saatsi, J. (Eds.), 2006. The Structural Foundations of Quantum Gravity. Clarendon Press.

CHAPTER 15

Mechanisms of Unification in Kaluza–Klein Theory

Ioan Muntean*

Abstract

In this chapter I discuss the attempts by Theodor Kaluza [Kaluza, T., 1921. Zum Unitätproblem der Physik. Sitzungsber. der K. Ak. der Wiss. zu Berlin, 966–972] and by Oskar Klein [Klein, O., 1926a. Quantentheorie und fünfdimensionale Relativitätstheorie. Zeitschrift für Physik 37 (12), 895–906; Klein, O., 1926b. The atomicity of electricity as a quantum theory law. Nature 118, 516], respectively, to unify electromagnetism and general relativity within a five-dimensional Riemannian manifold. I critically compare Kaluza's results to Klein's. Klein's theory possesses more explanatory power and unificatory strength and uses less types of brute facts than Kaluza's. The characteristic feature of Klein's theory is that it relies on an extrinsic element of unification, i.e. the wavefunction behavior, which is not intrinsic to **EM** or **GR**. Finally, I compare and discuss Kaluza's and Klein's theories in the context of Tim Maudlin's [Maudlin, T., 1996. On the unification of physics. Journal of Philosophy 93 (3), 129–144] ranking of unification and I clarify in what sense they constitute counterexamples to some of Margaret Morrison's [Morrison, M., 2000. Unifying Scientific Theories: Physical Concepts and Mathematical Structures. Cambridge University Press] assertions about unification.

Philosophers of science have discussed a great number of different cases of scientific unification.[1] However, the notorious forerunner to many unificatory attempts in string theory, the Kaluza–Klein theory, is barely mentioned as a peculiar case of unification in the philosophical literature.[2] I claim that the "Kaluza–Klein

* Department of Philosophy, University of California, San Diego, USA

1 The most comprehensive analysis of scientific unification is Morrison (2000) which contains an impressive number of illustrations. In the last years the practice of scientific unification has been discussed in Plutynski (2005), Ducheyne (2005), van Dongen (2002b). My analysis is not a general approach to scientific unification, but a study of the above analyses of the 'practice of unification' from which a limited number of general claims can be drawn.

2 Aitchison (1991), van Dongen (2002a), Weingard (nd, 1984) are among the few who discussed the Kaluza–Klein unification. In my view, a philosophical analysis of unification in string theory for example should originate in the discussion of Kaluza–Klein. See for example Weingard (nd) who explains why Kaluza–Klein is a special case of unification.

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00015-6
© Elsevier BV
All rights reserved

geometrization" of physical fields is a distinctive kind of unification whose study offers insights into the relationship between unification and explanation. More precisely, I want to answer the following questions about unification in Kaluza's theory and Klien's theory:

(I) What is specific to Kaluza–Klein unification and what does it teach us about unification in general?
II) Are unificatory mathematical structures in Kaluza–Klein equipped with explanatory power?
(III) Where should we place Kaluza's and Klein's cases among other gauge unifications?
(IV) What kind of brute facts do Kaluza and Klein rely upon?
(V) In what sense is Klein's unification better than Kaluza's?
(VI) What are the limitations of the Kaluza–Klein unification?

In order to clarify where my case study stands with respect to the existing literature, I depict in the first section the philosophical approaches to scientific unification relevant to my case study. In the second section I give a general description of the unification by "geometrization". In the third and fourth sections I describe in greater detail the steps toward unification taken by Kaluza and Klein, respectively. In the last two sections I directly address the above questions by discussing the novel element of unification in Kaluza and, respectively, in Klien.

1. PUZZLES OF SCIENTIFIC UNIFICATION

Unification is a universally agreed upon virtue of a theory, but at the same time it is a vague philosophical concept. Philosophers and scientists likewise struggle to define it, to rank known cases or at least to describe or deal with some of its aspects. Despite many efforts, scientific unification remains a conundrum.[3] It is vague in the sense that there is no general definition or criterion available; when defined, it is often vulnerable to charges of triviality, spuriousness or *ad-hocness*.[4] Examples of trivial or spurious unifications are often provided in the literature: unification consisting in the mere conjunction of child psychology and fluid dynamics is for example trivial, whereas a conjunction of Kepler's law and Boyle's law is spurious.[5] Feynman's clever example[6] in which all laws have the form $A_i = 0$ (for example $(F - ma)^2 = A_1$, $(F - G\frac{m_1 m_2}{r^2})^2 = A_2, \ldots$, and "the theory of everything" $\sum_i A_i = 0$, is frequently quoted against unification *tout court*. In these mock cases, the 'unificatory' theory makes no contribution (explanatory, confirmatory, interpretation of free parameters, etc.) in addition to what was already contained in the original

[3] Some would say that we have an "intuition" of it like "you know it when you see it". Looking at intuitively "borderline" cases surely helps, but this does not suffice. P. Teller expresses this uncertainty in a concise way: "I agree that unifications [and reductions] show something important about how our theories bear on the world. But I take the worries to show that we are very far from understanding what that 'something' is." (Teller, 2004, 443).

[4] My main target is to show that Klein improved significantly upon Kaluza's theory. I talk about *relative* spuriousness, ad-hocness and strength of unification.

[5] Maudlin (1996, 131), Kitcher (1981, 526).

[6] Feynman et al. (1993, 25-10-11).

theories. A derivation of a law from such a conjunction is clearly a pointless "self-explanation" or "self-confirmation". However, a general criterion for what makes a unification a compelling one, or what makes it trivial or spurious is not available.

Even when unification is not trivial, it may or may not be relevant or related to major topics in philosophy of science such as the realism/antirealism debate, confirmation, intertheoretical reductionism, causation, etc. Moreover, even if unification is not trivial and it is allegedly relevant to some other more stringent commitments like realism or empirical confirmation, the price to achieve it in some cases is nevertheless very high (the most notable case discussed by philosophers is the electroweak unification).

In this chapter I weasel out of providing general answers to all the questions related to these issues because I accept that unification is too vague a concept to qualify for a comprehensive approach at a general level. I prefer a more "pluralistic" talk about degrees of unification, stages or levels of unification or successful or unsuccessful unifications. Nonetheless, my case study reveals some unexpected aspects of the mechanism of unification. Among these, it illustrates the connection between unification and explanation.

(A) The strengths and weaknesses of unification qua *explanation*. A tentative definition of unification, inspired by the D-N model, would look like this: a unificatory theory (T_0) describes a set of phenomena previously described by two different theories (T_1 and T_2) by using fewer sentences (or "covering" laws). In the mid 20th century, the unity of science had been thought to operate in a reductionistic way: a unifying theory T_0, more general and more abstract than the old theories taken together, reduces T_1 and T_2. But even in the field of theoretical physics unification cannot be confined to reduction, as anti-reductionism and unification can coexist.[7] If not reduction, then what is the key concept of unification? In the heyday of the D-N model, it was suggested that the aim of explanation was *unification* i.e., "the comprehending of a maximum of facts in terms of a minimum of theoretical concepts and assumptions".[8] In Friedman's view, phenomena are explained if represented by a minimum of law-like sentences. "I claim that this is the crucial property of scientific theories we are looking for; this is the essence of scientific explanation—science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given."[9] T_0 proceeds by providing fewer types of brute facts (or independent phenomena) than T_1 and T_2 do. In explaining different results with the same theory, we reduce the number of brute facts we need and we unify our knowledge of the world. Unificatory theories are simpler (and maybe more beautiful); they increase our understanding of the world using less brute facts.

My case study will shed light on some difficulties with Friedman's account. Firstly, positing in an *a priori* way brute facts or trying to reduce their number by pseudo-explanations are both signs of weakness of a theory. What is important for

[7] Crystallography and solid-state physics "emerged" from the quantum theory without being reducible to it. Also, **GR**, which provides a unification of spacetime and the Newtonian gravitational field, is not a reduction, as neither of them survived "unscathed" in **GR** (Maudlin, 1996, 133).

[8] Feigl, Kneale and Hempel expressed similar views.

[9] Friedman (1974, 15).

a theory is not the sheer number of brute facts, but to get the *right* sorts of facts as brute.[10] Secondly, giving the difficulties of counting such brute facts, Ph. Kitcher suggested that reducing the number of *types* of facts generally is a better choice: "Science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same patterns of derivations again and again and, in demonstrating this, it teaches us how to reduce the number of types of facts we have to accept as ultimate (or brute)".[11] Thirdly, aside of explaining the world with less types of brute facts, the unificatory theory T_0 has to act as a "problem solver" for T_1 and T_2 without generating it own baggage of troubles. All these issues will be illustrated in the case study to follow.

The connection between unification and explanatory power of theories has been questioned in the last decade. Most notably, Margaret Morrison claimed that unification and explanation are "decoupled".[12] Rather than being a special case of explanatory power, unification is independent of explanation such that "they have little to do with each other and in many cases are actually at odds."[13] Using examples of unified theories, Morrison argued that "the mechanisms crucial to the *unifying* process often supply little or no theoretical *explanation* of the physical dynamics of the unified theory."[14] Many of her case studies against unification *qua* explanation are theoretical physics examples in which unity is usually understood in terms of derivability from a mathematical structure. The mathematical structure (for example the tensor calculus in Special Relativity), bestows scientific theories with a high level of generality making it applicable in a variety of contexts and suited to unifying different domains. However, for Morrison this unificatory mechanism of quantitative laws does not provide any explanation of the "machinery" or the mechanism of the phenomena.[15] The mark of a *truly* unified theory is "a specific mechanism or theoretical quantity/parameter that is not present in a simple conjunction, a parameter that represents the theory's ability to reduce, identify or synthesize two or more processes within the confines of a single theoretical framework".[16] Thus, Maxwell used a "substantial identification" of the optical aether with the electric ether on the base of the numerical identification of their velocity of transmission,[17] although the real unificatory element in Maxwell was the "displacement current". Even if such a factor is present, other troubles linger for unification. Weinberg's current in the electroweak unification is the parameter that unifies the parameters of the electromagnetic theory and those of the weak interaction, but for Morrison it has no explanatory power (here the Higgs mechanism explains the phenomena) and it is as arbitrary as the previous ones.[18] Morrison argues that in this case (as well as in **SR**, and to some extent in the case of

[10] Lange (2002, 99).
[11] Kitcher (1989, 432).
[12] See especially Morrison (2000), but also Morrison (1995, 1992).
[13] Morrison (2000, 1-2) and Morrison (2000, 64).
[14] Morrison (2000, 4).
[15] "The machinery is what gives us the mechanism that explains why, but more importantly *how a* certain process takes place." Morrison (2000, 3). One example of "machinery" quoted in Morrison (2000, Ch. 3) is Maxwell's explanation of the electrodynamics in terms of ether.
[16] Morrison (2000, 64).
[17] Morrison (2000, 98).
[18] Morrison (2000, 139).

the synthesis of Mendelian and Darwinian theories in biology) we have a *suspect* unity. Accordingly, many unification cases are less exemplary than believed and the mathematical structure alone does not imply true unification.

(B) Unification in theoretical physics. Physics is replete with claimed instances of unification: in seeking new theories not yet empirically confirmed, physicists often espouse a desire for theoretical virtues like unification and strive to reach it for reasons ranging from aesthetic considerations like simplicity and harmony, to more pragmatic reasons like the paucity of language or computability restrictions.[19] Morrison's conclusion raises a question about unification in physics: how explanatory is a unificatory theory? Philosophers of physics prefer to directly relate unification not to explanation, but to the way in which different forces can be captured within one and the same mathematical formalism. It is not uncommon to relate unification in physics not to explanation, but to the gauge symmetries and group of the unified theory T_0. There are in fact two major goals of unification of classical fields: unifying different force fields and, second, unifying a force field with its source.[20] In **SR** the first goal can be achieved by identifying the electric and magnetic fields with components of the tensor field $F_{\mu\nu}$ such that a Lorentz transformation transforms the components of one into the other. The distinction between electric and magnetic fields disappears in relativistic electrodynamics: electric and magnetic fields are eliminated from the ontology by being replaced by the field tensor which is frame-independent. We will see that such a mechanism of unification is only *partially* present in Kaluza–Klein theory.

In an attempt to rank the varieties of unification in theoretical physics, Tim Maudlin imposed three conditions on any non-trivial unification of two theories (T_1 and T_2): (a) T_1 and T_2 have to be consistent, (b) the field force in T_1 has to obey the same dynamics as the field force in T_2 and (c) there is a lawful (or nomic) correlation among the forces described by T_1 and T_2. The necessary conditions (a)–(c) constitute "a lower bound" of unification and I will discuss in detail whether Kaluza and Klein theories obey (a)–(c). At the other end of the spectrum Maudlin situated two cases of "perfect unification": in the theory of electrodynamics unification, as well as the unification of inertial and gravitational masses in **GR**. "Perfect unifications" provide novel predictions, too: for example, **GR** provided predictions that have been confirmed only much later. A perfect unification can play a role in arguments for realism: e.g., for believing that the entities postulated by **GR** are real.

In the case of theories describing various interactions among particles, unificatory theories can be ranked by appealing to their gauge symmetries.[21] Maudlin noticed that many gauge theories, praised as embodying unification, do not qualify as 'perfect'. For example, a trivial case of gauge unification is when two gauge theories T_1 with the symmetry group G_1 and neutral particle X_1 and, respectively, T_2 with G_2 and neutral particle X_2 are "pasted" into a product group $G_1 \otimes G_2$ without any further ado. The standard model itself was build up as the product group:

[19] The so-called **GUT** (Grand Unified Theories), the standard model and string theory are examples of theories having unification as a primary motivation.

[20] Weingard (nd, 1).

[21] Maudlin (1996), O'Raifeartaigh and Straumann (2000).

$SU(3) \otimes SU(2) \otimes U(1)$. These 'pasting' unifications are nothing more than a conjunction of several laws of dynamics. A next level of unification can be achieved when the product gives rise to new observable forces and observable particles created from *mixing* the groups G_1 and G_2 by a "mixing angle" between X_1 and X_2. In the case of the electroweak unification, the group is $SU(2) \otimes U(1)$. Even at this level, some physicists (H. Georgi, K. Moriyasu) suspect "a partial unification, at best". Finally, the upper level of gauge unification is premised on the simple gauge group simple group (which is not decomposable in a product, as above). Ranking Kaluza and Klein among gauge symmetries is a difficult task because gravity is not a gauge theory in a trivial sense: particles do not couple to the gravitational field, but they exist in spacetime. Even if primarily Kaluza–Klein is not a theory of interaction among particles and even if the gauge classification does not apply in this case, I will explain the importance of pasting gravitation and electromagnetism together in a way that is invariant to coordinate transformations.

There is an another point I want to stress at this stage. The original Kaluza–Klein theory is a *false* theory (for reasons whose analysis would take us beyond the scope of the present analysis). It is not clear yet whether an updated and refined version of the Kaluza–Klein theory has any chance of being true.[22] What I want to analyze here is the mechanism behind this unification program, irrespective of whether the theories involved are true or not.

2. SIMILARITIES AND DISSIMILARITIES BETWEEN GR AND EM

In order to understand the Kaluza–Klein mechanism, we need to remember the historical context after the discovery of **GR**. Some formal similarities between gravitation and electromagnetism were apparent to Einstein, G. Nördstrom and H. Weyl: both **EM** fields and gravitation are described by Poisson equations. Einstein's field equation $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu}$ was intended to show how *the metric* $g_{\mu\nu}$ responds to the presence of energy and momentum of *matter* represented by $T_{\mu\nu}$.[23] Similarly, the inhomogeneous Maxwell equation: $\partial_\nu F^{\mu\nu} = \mu_0 J^\mu$, describes how electric and magnetic fields respond to J_μ (which encodes the charges and the currents). Notwithstanding many differences in the nature of gravitation and electromagnetism,[24] in both cases we encounter partial differential equations (**PDE**) for the fields, with matter or charges as sources codified in

[22] See details in Wesson (2006, 5). See for example the results mentioned on the webpage of the "Space-Time-Matter Consortium" at URL: http://astro.uwaterloo.ca/~wesson/.

[23] Einstein contemplated the possibility to turn the "wood" of $T^{\mu\nu}$ (the matter) into the "marble" of $G_{\mu\nu}$ (the spacetime). For him matter was a term that 'infested' the pure and clean structure of $G^{\mu\nu}$. As Kaku commented: "By analogy, think of a magnificent, gnarled tree growing in middle of a park. Architects have surrounded this grizzled tree with a plaza made of beautiful pieces of the purest marble. The architects have carefully assembled the marble pieces to resemble a dazzling floral pattern with vines and roots emanating from the tree. To paraphrase Mach's principle: The presence of the tree determines the pattern of the marble surrounding it. But Einstein hated this dichotomy between wood, which seemed to be ugly and complicated, and marble, which was simple and pure. His dream was to turn the tree into marble; he would have liked to have a plaza completely made of marble, with a beautiful, symmetrical marble statue of a tree at its center." Kaku (1994, 99).

[24] There are major differences between the two theories. Maxwell equations are linear, but Einstein field equation is not. For an excellent philosophical discussion about the role of dissimilarities with **EM** in the genesis of **GR** see Norton (1992).

the right hand terms:

$$\begin{pmatrix} \text{“variation of a field”} \\ \textbf{PDE} \text{ in:} \\ F_{\mu\nu}, g_{\mu\nu} \end{pmatrix} \begin{pmatrix} \text{“related to”} \\ = \end{pmatrix} \begin{pmatrix} \text{“a measure of a source”} \\ \text{linear terms in:} \\ T_{\mu\nu}, J_{\mu}, \text{ etc.} \end{pmatrix}$$

In the light of these similarities, both fields g and F could stem from one and the same universal tensor and could have a common dynamics (i.e. the condition b) in Maudlin) so it becomes plausible to ask whether it is possible for the structure of spacetime to explain the **EM** equations, as it does explain gravity. However, there is a difficulty here. **GR** had already been designed to include the **EM** field: all fields but g, as well as matter and charges, were present in the $T_{\mu\nu}$, so in this sense gravity was geometrized, while electromagnetic fields were not. In **GR** the gravitational field has become a geometrical object, while **EM** works with ordinary fields that are the effect of charged particles and currents. In its more general form, the “geometrization” program endeavored to express F as a feature of the geometry of spacetime. Some ideas towards this goal were already available in the 1920s: the **EM** field could perhaps be considered a part of the curvature, part of the connection or part of the metric. Kaluza and Klein explicitly preferred the latter option. Their “geometrization” program was intended to move all the non-material fields within g and to unify the fields by embedding them all into the geometry of spacetime. By this “geometrization”, the fields become aspects of the same entity, the metric tensor, such that geometry and physics are no longer distinct ways of describing the world.[25]

3. KALUZA'S UNIFICATION VIA THE FIFTH DIMENSION

Starting from Einstein's theory, Theodor Kaluza[26] tried to provide such a geometrical explanation for electromagnetism. Kaluza's approach was more speculative than computational or empirical as it aimed to remove the *duality* of gravity and electricity, “while not lessening the theory's [of gravity] enthralling beauty”.[27] Kaluza's starting point was the idea that if the universe is empty of matter and charges the only real entity is g. This is the “vacuum hypothesis” (no matter, no charges present):

$$\texttt{VACUUM:} \quad T_{\mu\nu} = 0$$

Giving VACUUM, where is the place for the **EM** field? The intuitive answer is: somewhere in the expression of g itself. But both theories have their own vacuum solutions, and if one tries to describe gravitation and electromagnetism in one scheme, a problem is that the presence of electromagnetic vacuum solutions makes it impossible to have a “gravitational vacuum”. The attempt to unify the two vacuum solutions fails for various reasons. One of them is that in 4-D there is no way to add the field tensor $F_{\mu\nu}$ to the Christoffel symbols $\Gamma^{\lambda}_{\mu\nu}$, to preserve

[25] Weingard (nd, 3).
[26] Kaluza (1921, 860).
[27] Kaluza (1921, 865).

their properties and to impose later on $R_{\mu\nu} = 0$. Christoffel symbols are defined only up to the first derivatives of a single field and they represent the "displacement" of a vector. Therefore, the "geometrization" of $F_{\mu\nu}$ is not possible in a *four-dimensional Riemannian* manifold. Unlike Weyl's unification, Kaluza kept the metric (pseudo-)Riemannian. What he changes is the *dimensionality* of the g, R and Γ tensors.[28]

The field equations in 5-D. By "calling a fifth dimension to the rescue",[29] Kaluza managed to express the **EM** field as part of the metric g. There is thus room for all fields within g_{mn} and only matter and electrical charges (if any) are present in T_{mn}. Kaluza added to the Riemannian $g_{\mu\nu}$ one row and one column:[30]

$$ds^2 = g^{(5)}_{mn}\, dx^m\, dx^n \tag{1}$$

All the expressions of tensors and the relations between them, as well as the Christoffel symbols, are simply generalized from four to five dimensions. Kaluza speculated about a formal similarity between the Christoffel symbols in 5-D and the 4-D expression for $g_{\mu\nu}$ and $F_{\mu\nu}$. The 4×4 part of $g^{(5)}_{mn}$ can simply be equal to the original $g_{\mu\nu}$. So where is the $F_{\mu\nu}$ to be placed? The simplest way is to divide $g^{(5)}$ in three sectors as follows:

$$g^{(5)}_{mn} = \left(\begin{array}{c|c} g_{\mu\nu} = \text{'G' sector} & g_{4\nu} = \text{'EM' sector} \\ \hline g_{\nu 4} = \text{'EM' sector} & g_{44} = \phi = ? \end{array}\right) \tag{2}$$

which can accommodate the $g_{\mu\nu}$ tensor in the 'G' sector as well as the A_μ vector in the 'EM' sector. More information about these sectors can be gathered from the Christoffel symbols:

$$-2\Gamma_{4\mu\nu} = \partial_4 g_{\mu\nu} + \partial_\mu g_{\nu 4} - \partial_\nu g_{4\mu} \tag{3}$$

$$-2\Gamma_{\mu\nu 4} = \partial_\mu g_{\nu 4} + \partial_\nu g_{4\mu} - \partial_4 g_{\mu\nu} \tag{4}$$

`CYLINDER CONDITION:` This is easy to see why there is a surplus structure in the 5-D metric and much of this has to be stripped away. Here is Kaluza's suggestion: in order to take in the homogeneous Maxwell equations: $\partial_\mu F_{\nu\kappa} + \partial_\nu F_{\kappa\mu} + \partial_\kappa F_{\mu\nu} = 0$, one term out of three is always set to zero in (3) and (4) such that $\Gamma_{4\mu\nu}$ and $\Gamma_{\mu\nu 4}$ will contain *only* **EM** terms. The best option is to hypothesize that $\partial_4 g_{\mu\nu}$ vanishes. This is formally the origin of the so called "cylinder" condition, arguably the very core of the Kaluza–Klein unification:

$$\texttt{CYL:} \quad \partial_4 g_{mn} = 0$$

[28] Parenthetically, we need to mention that as early as 1914, G. Nördstrom expressed the metric as a 5×5 matrix. His theory is less known than Kaluza's and has had only a slight impact on the scientific community. He added another spatial dimension to the four existing ones in order to obtain an Abelian five-vector gauge field for which a Maxwell-like equation can be written, including a conserved 5D current. He was the first to explicitly claim that "we are entitled to regard the four-dimensional space-time as a surface in a five-dimensional world." The major difference between Nördstrom and Kaluza is that the former found gravity by applying **EM** to the 5D world, whereas the latter applied **GR** to it Smolin (2006, 47). From my point of view, Kaluza scores better than Nördstrom in respect of unification.

[29] Kaluza (1921, 967).

[30] Latin indices are numbers from 0 to 4 and Greek indices are from 0 to 3; vectors or tensors with Latin indices are 5-dimensional. Here x^0 is the time coordinate. Time is the zeroth component of a 4-vector and $x^1 \ldots x^3$ are the Cartesian spatial coordinates.

We experience three dimensions of space and one of time because there are fields in these four 'directions' which are *not* constant. Null, or higher order variations of the fields on x^4, means that the world is "cylindrical": every point $P(x^0, \ldots, x^4)$ is indistinguishable from another point P' having the coordinates $P(x^0, \ldots, x^4 + \delta x^4)$ if g is constant in the fifth direction.[31] P and P' are still distinct, notwithstanding the values of all possible physical fields being equal or having very close values at these points. Consequently it is natural to call $F_{\mu\nu}$ a "degenerate" (*verstümmelte*) form of the 5-D Christoffel symbols and to reduce it to them:

$$\texttt{ID}_1\text{:} \quad \Gamma_{4\mu\nu} = \alpha F_{\mu\nu} \tag{5}$$

$$\Gamma_{\mu\nu 4} = -\alpha(\partial_\nu A_\mu + \partial_\mu A_\nu) \tag{6}$$

$$\Gamma_{44\mu} = \partial_\mu \phi \tag{7}$$

where ϕ is an arbitrary scalar field, not yet interpreted.

WEAK FIELD: In order to provide analytical solutions to the field equations, one commonly assumes the perturbation formulation of **GR** in which the metric differs only a little from its Euclidean value $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ (where $\eta_{\mu\nu}$ is a Minkowskian metric and "the perturbation" h is taken such that $|h_{\mu\nu}| \ll 1$). This leads to "linearized gravity". In order to conduct his analysis, Kaluza assumed that the third and fourth terms in the Ricci curvature in 5-D:

$$R^m_{ijk} = \partial_j \Gamma^m_{ik} - \partial_k \Gamma^m_{ij} + \Gamma^n_{ik}\Gamma^m_{nj} - \Gamma^n_{ij}\Gamma^m_{nk} \tag{8}$$

are of the form Γ^2, and since Γ is of first-order, these contribute only to second order and can be discarded.

$$\texttt{WEAK:} \quad R^m_{ijk} \cong \partial_j \Gamma^m_{ik} - \partial_k \Gamma^m_{ij}.$$

The Ricci tensor obtains a simpler form, too:

$$R_{\mu\nu} = \partial_\lambda \Gamma^\lambda_{\mu\nu} \tag{9}$$

$$R_{4\nu} = \alpha \partial_\lambda F^\lambda_\nu \tag{10}$$

$$R_{44} = -\partial_\mu \partial^\mu \phi \tag{11}$$

Kaluza supposed that the 5-D world is empty, so both the Ricci scalar and the curvature tensor vanish:

$$R_{mn} = 0 \quad \text{and} \quad R = 0 \tag{12}$$

Then, what does the 5-D vacuum generate? The assumptions that have been made yield an unexpected number of direct results, including the derivation of vacuum solutions in 4-D from the 5-D vacuum solution. By mimicking some of the **GR** techniques, Kaluza was able to infer the following equations:

- The 5D metric:

$$g^{(5)}_{mn} = \begin{pmatrix} g_{\mu\nu} & 2\alpha A_\mu \\ 2\alpha A_\nu & 2\phi \end{pmatrix} \tag{13}$$

[31] This analogy is from Einstein and Bergmann (1938).

- Homogeneous Maxwell equations from Christoffel symbols, (5) and (6).
- Einstein field equations 4-D from (9).
- A Poisson-like equation for ϕ from (11).
- The components of the energy momentum tensor in 5-D. In the WEAK approximation, the Ricci scalar is of higher order in h and the Einstein equations in 5-D are:

$$R_{mn} = \kappa T_{mn} \tag{14}$$

Again, from (10) and the inhomogeneous Maxwell equation, one can identify the components of T_{mn} as:

$$\text{ID}_2: \quad T_{\mu 4} = J_\mu$$

so Kaluza has bordered the 4-D energy momentum tensor $T_{\mu\nu}$ with a vector representing the currents and densities of charges. It is easy to show that $T_{55} = 0$ and then $T_{\mu\nu}$ is:

$$T^{(5)}_{mn} = \left(\begin{array}{c|c} \text{matter and densities: } T_{\mu\nu} & J^\mu \\ \hline \text{currents and charges: } J_\mu = (c\rho\ j_1\ j_2\ j_3) & 0 \end{array}\right) \tag{15}$$

- Maxwell inhomogeneous equation from WEAK, CYL, (5) and (10).

Even if Kaluza thus accomplished the intended unification program, two major aspects of **GR**—the geodesics and the definition of energy have still to be explicitly analyzed.

Geodesics in 5-D. The ideal situation would be like this: a small, charged test particle in 5-D space moves along a geodesic in 5-D and its projection in 4-D is the expected trajectory of a particle with mass M and charge q in curved spacetime in which an electric field tensor $F_{\mu\nu}$ is present which is *not* a geodesic:

$$\frac{d^2x^\rho}{dt^2} + \Gamma^\rho_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt} = -\frac{q}{Mc}F^\rho_\mu\frac{dx^\mu}{dt} \tag{16}$$

Since exact calculations are extremely difficult, Kaluza assumed a "slow motion approximation" (commonly used in **GR**) in which the 5-velocities $U^m = \frac{dx^m}{ds}$ are such that $U^m \cong (1, \vec{0}, U^4)$ and $ds^2 \cong d\tau^2$, where τ is the proper time. In this case $T^{mn} = \mu_0 U^m U^n$ and in order to estimate the geodesics, terms (U^4) are needed. By generalizing the 4-D geodesic equation (parameterized by λ) and by employing (9)–(11) a general equation of motion can be inferred:

$$\frac{d^2x^m}{d\lambda^2} + \Gamma^m_{ab}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda} = -\sqrt{2\kappa}\,F^m_n\frac{dx^n}{d\lambda}\frac{dx^4}{d\lambda} - \partial^m\phi\frac{dx^4}{d\lambda}\frac{dx^4}{d\lambda} \tag{17}$$

In order to identify this with (16), Kaluza chose the parametrization such that $\lambda = \tau \cong t$ and supposed that $U^4 = \frac{dx^4}{d\lambda}$ for macroscopic particles is small such that the last term in (17) vanishes. The third identification is:

$$\text{ID}_3: \quad U^4 = \frac{dx^4}{dt} = \frac{q}{Mc\sqrt{2\kappa}}$$

The interpretation of $\mathtt{ID}_3$ can raise difficulties but it also constitutes a powerful tool for explaining electromagnetism. Two particles in 4-D which have the same mass and the same initial conditions and differ only in respect of their charge will follow two trajectories which are both projections of a geodesic in 5-D, because they have different U^4.[32] Had we started with the small velocity approximation, we would want U^4 to be close to zero. The formalism applies only to relatively small velocities *and* to charges of $\rho_0/\mu_0 \ll 1$, which seems kosher for all practical purposes. But this second approximation is unsatisfactory for atomic dimensions where U^4 is not at all small for electrons or protons. In this case, the slow motion is no longer met and the motion of an electron is *not* a geodesics in $\mathbb{R}^5$ as U^4 is enormously large. This means that Kaluza's theory would not work for subatomic particles which is the major drawback of his theory.

4. OSKAR KLEIN'S COMPACTIFICATION OF x^4

Five years after Kaluza's paper was published, Oskar Klein wrote a paper and a note in *Nature* in which he dealt with the idea of unification of **EM** and **GR** by analyzing not only the $g^{(5)}$ field, but also the wavefunction on a 5-D manifold.[33] The first part of the paper was inspired by Kaluza and his treatment of the $g^{(5)}$ field, although the legend has it that Klein carefully read Kaluza only after he had finished writing his paper.[34] He himself started from the aforementioned similarities between **GR** and **EM**,[35] and postulated in 5-D the Riemannian metric (1), the forms of Ricci tensors and Christoffel symbols from **GR**. Klein assumed that the 15 quantities of the symmetric tensor g_{mn} would accommodate the 10 independent components of $g_{\mu\nu}$ plus the four components of A_μ. In order to fit these into g_{mn} and by echoing Kaluza's $\mathtt{CYL}$, Klein imposed some conditions on the coordinate system of the 5-D space:

- The first four coordinates are identical to the ordinary spacetime coordinates;
- The cylinder condition ($\mathtt{CYL}$): the fields do not depend on x^4;
- $g_{44} = a$, where a is a constant.

The first three are present under various guises in Kaluza. The latter is new and in Klein, it becomes central. It is worth mentioning that $\mathtt{CYL}$ is just a working hypothesis: later in this paper and in the note to *Nature*, Klein would replace it with the compactification ($\mathtt{COMP}$).[36] It can be proven that the only infinitesimal

[32] I'll offer a more comprehensive discussion of this issue in Section 4.

[33] Klein (1926b, 1926a). In Klein (1928) he came back to the problem of the unification and restated the main idea of compactification in direct relation to conservation laws.

[34] In his autobiographical note Klein recalls: "When Pauli came to Copenhagen [in 1925], I showed him my manuscript on five-dimensional theory and after reading it he told me that Kaluza some years before had published a similar idea in a paper I had missed. So I looked it up [...] I read it rather carelessly but quoted, of course, in the paper I then wrote in a spirit of resignation. [...] In the paper I tried, however, to rescue what I could from the shipwreck." (Ekspong, 1991, 111).

[35] Witness Klein's confession again: "The similarity struck me between the ways the electromagnetic potentials and the Einstein gravitational potentials enter the [relativistic Hamilton–Jacobi equation for an electric particle], the electric charge in appropriate units appearing as the analogue to a [fifth] momentum component, the whole looking like a wave front equation in a space of [five] dimensions. This led me into a whirlpool of speculation, from which I did not detach myself for several years and which still has a certain attraction for me." (Klein recollecting in 1989 the early 20s) (Ekspong, 1991, 108).

[36] I will come back on this issue later (p. 289sqq.)

coordinate transformation which satisfies these conditions is:[37]

$$x^\mu \to x^\mu + \xi^\mu(x^\nu) \tag{18}$$

where ξ are smooth functions of only the first four coordinates $x^0 \ldots x^3$. For such a transformation, the only metric tensor that preserves the line element ds^2 (see (1)) has the form:

$$g^{(5)}_{mn} = \begin{pmatrix} g_{\mu\nu} + \mathbf{A}^{(5)}_\mu \mathbf{A}^{(5)}_\nu & \mathbf{A}^{(5)}_\mu \\ \mathbf{A}^{(5)}_\nu & 1 \end{pmatrix} \tag{19}$$

where $\mathbf{A}^{(5)}$ is a 5-vector of which all first four components transform[38] like the covariant components of the **EM** field and $\mathbf{A}^{(5)}_5 = 1$. The simplest way is to *identify* again the four components of this 5-D vector with the **EM** vector potential A_μ:

$$\texttt{ID}_4\text{:} \quad \mathbf{A}^{(5)}_\mu = A_\mu \tag{20}$$

The constant field ϕ is plugged into the expression of the metric in order to replace the $g_{44} = 1$:[39]

$$g^{(5)}_{mn} = \begin{pmatrix} g_{\mu\nu} + \phi A_\mu A_\nu & \phi A_\mu \\ \phi A_\nu & \phi \end{pmatrix} \quad \text{with } \phi = \text{const} \tag{21}$$

(A) Klein's metric. Despite these similarities, there are some important differences between Klein's and Kaluza's assumptions regarding the topology of the fifth dimension. Klein's metric is (as before, ϕ is taken constant):

$$ds^2 = (g_{\mu\nu} + A_\mu A_\nu)\, dx^\mu\, dx^\nu + 2\phi A_\mu\, dx^\mu\, dx^4 + \phi A_\mu (dx^\mu)^2 + \phi\, dx^4\, dx^4 \tag{22}$$

In order to show that $\texttt{ID}_4$ is not arbitrary, Klein inferred Einstein's field equation and Maxwell equations from a variational principle (instead of guessing an expression for the Ricci tensor like Kaluza did) by requiring the minimization of the Hilbert action under the variation of the metric $\delta g^{(5)}_{mn}$ and of its first derivative $\partial_l g^{(5)}_{mn}$:

$$S_H = \int \mathcal{L}_1\, d^5x \tag{23}$$

where $\mathcal{L}_1 = R^{(5)}\sqrt{-g^{(5)}}$ is a Lagrange density of fields and $R^{(5)}$ is a Ricci-like invariant scalar. By accepting Kaluza's WEAK, Klein disregarded the contribution of the last two terms in (8) and proceeded by applying the CYL. As $R^{(5)}\sqrt{-g^{(5)}}$ does

[37] See Klein (1926a, 896), but the hereby terminology is from Bergmann (1942).

[38] The vector **A** is a vector field employed in projective geometry, see Bergmann (1942, 274).

[39] Although Klein did not provide a matrix form of the metric, I use a matrix representation here. I do not adopt the exponentiation of ϕ from Duff (1994). In the "projective geometry" formulation of Veblen and Hoffmann the metric suffers an extra coordinate transformation $x^4 \to e^{x^4}$. The importance of the scalar field ϕ will be discussed later. See for details O'Raifeartaigh and Straumann (2000), Bergmann (1942), van Dongen (2002a). Witness the presence of the $\mathbf{A}^{(5)}$ in the 4×4 part of $g^{(5)}$.

not depend on x^4, the integral in (23) splits into two integrals like:

$$\int dx^4 \int R^{(5)} \sqrt{-|g^{(5)}|}\, d^\mu x$$

The action is an integral in 4-D only:

$$S = -\int d^4x \sqrt{-g} \left(\frac{R}{\kappa^2} + \frac{1}{4}\phi F_{\mu\nu}F^{\mu\nu} + \frac{1}{6\kappa^2} \frac{\partial^\mu \phi\, \partial_\mu \phi}{\phi^2} \right) \tag{24}$$

The first integral in (24) is simply the action for gravity in 4-D,[40] while the second is an action of the electromagnetic field of a stress-energy tensor given by Maxwell equations and the third is the Klein–Gordon equation of the scalar field ϕ.[41] By minimizing the action $\delta S_H = 0$, the result is a system of two equations:

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = \kappa T^{\mu\nu} \tag{25}$$

$$\partial_m \sqrt{-|g|} F^{\mu m} = 0 \tag{26}$$

This result is strikingly close to Kaluza's. Through a minimization of the action of the $g^{(5)}$ field in 5-D, Klein recovered the gravitation field of Einstein field equation and both Maxwell equations for vacuum.

(B) Charges and matter on geodesics. The first good news for Klein was that the metric (21) yields the right form of the geodesics in 5-D.[42] Indeed, Klein added to the action (23) a Lagrange density for the motion of n free charged particles. The total Lagrange density in the presence of fields and n probe particles then is:

$$\mathcal{L} = \mathcal{L}_1 + \sqrt{-g^{(5)}}\kappa \sum_{i=1}^{n} g^{mn} \frac{dx_i^m}{d\lambda} \frac{dx_i^n}{d\lambda}$$

Similar to Kaluza's $\mathtt{ID}_3$, in order to derive the geodesics in 5-D, Klein interpreted the velocity on the fifth axis as proportional to the charge of the particle:

$$U^4 = \frac{e}{c} \frac{1}{\frac{d\tau}{d\lambda}} \tag{27}$$

where as usual $d\tau = \frac{1}{c}\sqrt{-ds^2}$ is the proper time in 5-D, λ is a parameter of the geodesics and e is the electrical charge of the electron. From (22) and the Ricci tensor, Klein inferred the 5-D geodesics.[43] On such geodesics, the Lagrange function $L = \frac{1}{2}(\frac{ds}{d\tau})^2$ provides the definition of the 5-D momentum:

$$p_i = \frac{\partial L}{\partial(\frac{dx^i}{d\lambda})} \tag{28}$$

[40] O'Raifeartaigh and Straumann (2000, 9).

[41] Overduin and Wesson (1998, 15).

[42] "I became immediately very eager… to find out whether the Maxwell equations for the electromagnetic field together with Einstein's gravitational equations would fit into a formalism of five-dimensional Riemann geometry (corresponding to four space dimensions plus time) like the four-dimensional formalism of Einstein. It did not take me a long time to prove this in the linear approximation, assuming a five-equation, according to which an electric particle describes a five-dimensional geodesic." (Ekspong, 1991, 109–110).

[43] Klein (1926a, 899).

As there is no explicit dependence of L on x^4, we will always have a constant momentum on the fifth axis. The calculations render for an electron:

$$p_4 = \frac{e\sqrt{a}}{c\sqrt{2\kappa}} \tag{29}$$

where a is the constant value of the scalar field ϕ, so *if* the field ϕ is kept constant, p_4 has the same value at any point of spacetime.

(C) The 5-D wavefunction. The second part of the 1926 paper and the note in *Nature* are directly related to two major developments of both relativity and quantum mechanics. Here Klein inferred for the first time the form of the relativistic wavefunction for a spinless particle.[44] Klein endeavored to connect quantum results with the analysis of geodesics in 5-D. Instead of describing only particles on the manifold, Klein explicitly relied on de Broglie's treatment of quantum phenomena by analogy with mechanics.[45] Klein studied the differential form of a "ray" of a wave and then tried to identify it with the equation of the trajectory of a charged particle. The central point of the wave-particle analogy of de Broglie is the definition of the momentum operator by the operator "nabla" $\hat{\mathbf{p}} = -i\hbar\nabla$:[46]

$$\hat{P}_m = \frac{\partial}{\partial x^m} \tag{30}$$

Klein took a generalized form of a wave in 5-D:

$$\Box_g \psi = a_{ij}\left(\frac{\partial^2}{\partial x^i \partial x^j} - \Gamma^k_{ij}\frac{\partial}{\partial x^k}\right)\Psi = 0 \tag{31}$$

where $\Box_g$ is a wave operator in 5D, a_{ij} are some functions of the coordinates only and $\Psi = \Psi_0 e^{i\omega\Phi(x^m)}$ is a harmonic wave in 5-D.[47] Klein analyzed the solutions in two cases.[48] For ω large enough, the wave operator will have terms only in ω^2. The remainder is an equation of the phase ϕ: $a_{ik}(\frac{\partial\phi}{\partial x^i}\frac{\partial\phi}{\partial x^j}) = 0$ and the Hamiltonian of the propagation of the wave can be written as: $H = \frac{1}{2}a_{ik}P_iP_j = 0$, which is similar to the one in the Hamilton–Jacobi formalism. Rays are geodesics of the differential form: $a_{ik}dx^i dx^k = 0$ and the equation of motion of a charged particles is: $\mathcal{L} = \frac{1}{2}\frac{d\theta}{d\lambda} + \frac{ds}{d\lambda}$. In accordance to the duality postulated by de Broglie, the particle

[44] This equation was published in the same year by Klein, V. Fock and Gordon (allegedly Schrödinger had first discovered and immediately rejected it in 1925 because it could not explain spin). Klein's manuscript was submitted to the editors of *Zeitschrift für Physik* in April 1926, whereas Fock's and Gordon in July, respectively in September. Fock (1926) also used a 5-D formalism, very similar to Klein's. Not much attention has been paid to the fact that the Klein–Gordon equation originated in an explicit 5-D formalism.

[45] "[I tried] to learn as much as possible from Schrödinger and also from de Broglie, whose beautiful group velocity consideration impressed me very much even if by and by I saw that it did not essentially differ from my own way by means of the Hamilton–Jacobi equation. From Schrödinger I learnt in the first place his definition of the non-relativistic expressions for the current-density vector, which it was then easy to generalize to that belonging to the general relativistic wave equation. In this, after Schrödinger's success with the hydrogen atom, I definitely made up my mind to drop the possible non-linear terms, although I was still far from certain that this was more than a linear approximation. Also I derived the energy-momentum components, which in the five-dimensional formalism belonged to the current-density vector. These I published much later, due to the appearance in the meantime of a paper by Schrödinger containing the corresponding non-relativistic expressions." (Ekspong, 1991, 111–112).

[46] van Dongen (2002a, 5).

[47] Klein (1926a, 900).

[48] I do not discuss here the case of small ω in which the Klein–Gordon equation originated.

is represented by the wave so the rays coincide with the particle's trajectory.[49] The results are: $p_i = \frac{\partial \mathcal{L}}{\partial \frac{dx^i}{d\lambda}}$ and respectively $p_4 = \beta(\pm\frac{e}{c})$ where β is a constant. Because of $\Phi = -x^4 + S(x^0, x^1, x^2, x^3)$, the wavefunction Ψ can be separated into: $\Psi = \exp(i\omega x^4)\Psi(x^\mu)$. The conservation of phase along a closed trajectory in the fifth dimension is:

$$\omega \oint p_4 \, dx^4 = 2\pi n \tag{32}$$

and as the Hamiltonian of this wave is zero, the phase is conserved.

(D) Compactification on x^4 and the new argument restated. In *Nature*, Klein proposed a major turnover. "The charge q, so far as our knowledge goes, is always a multiple of the electronic charge e, so that we may write $p_4 = n\frac{e}{k}$ with $n \in \mathbb{Z}$. This formula suggests that the atomicity of electricity may be interpreted as a quantum theory law."[50] He hinted at the idea that the momentum along x^4 is always quantized.[51] Though it is not simply a classical "quantity of motion", quantum mechanical momentum has some properties of classical mechanical momentum (associated to moving particles or to waves). But the quantum momentum sometimes has a discrete spectrum, i.e. it is *quantized*. Because p^4 in (29) depends linearly on e, which is quantized, one may ask whether it is quantized, too. In polar coordinates, $\dot{\phi}$ or $\dot{\theta}$ are velocity-like quantities (they are actually angular velocities and there is an "angular momentum"), whereas p^4 is different. The analogy used by Klein has a purely heuristic role, and was inspired by early quantum results on closed orbits. The mathematical structure in both cases is that of a periodic function, *ergo* the idea of a Fourier expansion (used only later). However, while the hydrogen atom can be represented in coordinates in which $\phi = \phi + 2n\pi$, the atom itself does not live in a compactified space. The analysis of the wave in 5-D provided the idea of *compactification*. By taking into account de Broglie's hypothesis, one can infer:

$$p_4 = ne/c\sqrt{2\kappa} = n\hbar/\lambda_4 \tag{33}$$

where λ_4 is the radius of the closed circle on x^4. If one knows the quanta of electrical charge, from (33) one can deduce the compactification factor $\lambda_4 = 0.8 \cdot 10^{-30}$ cm. Klein identified geometrically the points P and P' separated by $2\pi\lambda_4$ and rejected the linear geometry of x^4 by the compactification hypothesis:

COMP: The x^4 axis is closed with a period of λ_4.

The new form of Klein's argument, the one usually cited, is obtained by replacing CYL with COMP. Instead of postulating the same values for the physical fields on

[49] In 1924 de Broglie's associated to each bit of energy with mass m_0 a periodic wave with a wavelength: $\nu_0 = m_0c^2/h$ de Broglie (1924, 11). The group velocity of this wave is the same as the velocity of the mass. Sommerfeld's condition for stability on hydrogen orbit can be inferred as conservation of phase. Schrödinger had anticipated in 1922 de Broglie's result that Weyl's scale factor (the exponential factor ϕ that relates the lengths of a rod parallel transported from P to P': $l_{P'} = l_P \exp \int \phi_i \, dx^i$) for closed orbits was an integral power of some universal constant. See Schrödinger (1923), Vizgin (1994). Klein took inspiration from de Broglie's thesis.

[50] Klein (1926b).

[51] Wave mechanics provided Klein with a clear form of a momentum on the fifth axis. But moving along x^4 is not simply a mechanical change of coordinates. This can be troublesome because it was for the first time when momentum had a non-dynamical interaction.

x^4, Klein took a different, geometrical stance: he supposed that the axis is curled with a very small radius. The consequence of the initial argument (COMP) was promoted to a hypothesis of the new argument and the hypothesis of the old argument (the quantization of charge) became a consequence of the new one. The new hypothesis COMP is then used to infer the quantization of charge and the new symmetry group of the theory. The smallness of λ_4, which is less than the Planck length, is the only reason why extensions on x^4 cannot be observed by macroscopic observers. Klein realizes that the discreteness of the charge spectrum, via the de Broglie relation, leads to a discrete wavelength in the fifth direction. In the new argument, given the value of λ_4, COMP explains CYL. The coordinate transformation allowed by COMP is $x^4 \rightarrow x^4 + \xi^4(x^\nu)$ (see (18)). Two points P and P′ are *identical* iff $x'^4 = x^4 + 2\pi\lambda_4$. As COMP is not a coordinate variant of the theory, the new structure of x^4 is not a mere alternative *representation*, but it reflects the structure of x^4 (unlike for example the case of polar coordinates, there are no transformations that remove the symmetry $S(1)$ and linearize x^4). Klein went well beyond the periodicity of coordinates by stipulating COMP. If two particles have the same initial condition in 4-D x_0^μ but different ratios q/M, they will fall under the *same* geometrical treatment in 5-D since there they follow geodesics. Obviously, this is an improvement over Kaluza's approach. Klein's metric does not need the small velocity approximation used by Kaluza and solves the problem of geodesics. From this we can infer the quantization of the charged particle as being imposed by (32). This means that if the fifth dimension is compactified with a period of $2\pi\lambda_4$, then the electrical charge appears *quantized* in 4-D.

(E) Only the first mode of Fourier expansion of fields is relevant. In Klein's days the fields $g_{\mu\nu}(x)$, $A_\mu(x)$ and $\phi(x)$ were thought to be mathematical objects which transform under four-dimensional general coordinate transformations. Klein did not notice that if COMP is assumed, then all fields are periodical on x^4 and consequently they can be Fourier expanded, having all other 4-D fields as coefficients.[52] The value of the 4-D field is reducible to an infinite number of values such that the first one is independent of x^4. For example, for g one can write:

$$\begin{aligned} g_{\mu\nu} &= \sum_{n=0}^{n=\infty} g_{\mu\nu}^{(n)}\left(x^0, x^\mu\right) e^{inx^4/\lambda_4} \\ &= g_{\mu\nu}^{(0)}\left(x^\mu\right) + g_{\mu\nu}^{(1)}\left(x^\mu\right) e^{ix^4/\lambda_4} + g_{\mu\nu}^{(2)}\left(x^\mu\right) e^{i2x^4/\lambda_4} + \cdots \end{aligned} \tag{34}$$

and similarly for A_μ and ϕ. He assumed that given the smallness of λ_4, all terms with $n > 0$ are large enough to not be visible from our 4-D world. Consequently, only the *first term* ($n = 0$) counts. This means that there is a ambiguity between the "real" 5-D tensor (or vector or scalar) and its 4-D "representation" ($g(x^\mu)$). The expectation values of these fields: $\langle g_{\mu\nu}\rangle$, $\langle A_\mu\rangle$, $\langle\phi\rangle$ were interpreted decades later as masses of particles. As Duff remarks, in today's parlance, the Fourier coefficient

[52] Einstein used this expansion as early as 1927. It can be shown that Klein's wavefield in 5-D is equivalent to the Fourier expansion. See van Dongen (2002a, 190). But Klein's analysis of the wavefunction in the harmonic approximation is close enough to a Fourier expansion. Albeit not present in Klein's paper, the Fourier analysis can be performed within Klein's model.

of order zero describes a graviton (spin 2), a photon (spin 1) and a dilaton (spin 0). Indeed, the masslessness of graviton $\langle g_{\mu\nu} \rangle = 0$ is due to the general covariance of *GR* (which can be interpreted as a gauge invariance); the masslessness of photon $\langle A_{\mu\nu} \rangle = 0$ is due to the gauge invariance; the masslessness of the dilaton $\langle \phi \rangle = 0$ is due to it being a Goldstone boson.[53]

5. KALUZA'S UNIFICATION: ITS LIMITS AND ITS PROMISES

In this section I intend to connect Kaluza's theory to the literature on scientific unification and to answer the questions raised in Section 1.

There is no "real unificatory element" or "machinery" (such as the "displacement current" in Maxwell) in Kaluza. Instead of a "theoretical parameter", Kaluza depicts a mathematical operation that unifies. Similar to Maxwell's case, Kaluza used $\texttt{ID}_1$-$\texttt{ID}_3$ to explain why we have the illusion of **EM** and **GR** as disparate theories: with the help of some approximations, the IDs enabled Kaluza to represent the **EM** and **GR** interactions with one and the same formalism and to infer a geodesic equation; from $\texttt{ID}_1$ he inferred the form of the metric tensor g_{mn} and from $\texttt{ID}_2$, the geodesic equation for macroscopic objects; $\texttt{ID}_3$ had helped him to provide an interpretation for p^4. The IDs provide answers to "why" questions such as: Why is it apparent that **EM** phenomena are independent of gravitational phenomena? Why do macroscopic charged particles not move on geodesics in 4-D? Why do **GR** and **EM** obey Poisson equations? I conclude that question (II) can be answered in the affirmative because Kaluza provides explanatory power along with unification, *pace* Morrison's claim.

(A) The GR-EM coupling and Kaluza's unification. One would like to have a **SR**-type of unification where the "electric field" by itself and the "magnetic field" by itself were doomed to fade away. In the prototypical case of unification of electric and magnetic forces within **SR**, the theory proves that they are descriptions of one and the same physical entity. This is not the case with Kaluza's theory. He intended to provide the strongest unification possible, but his formalism is not powerful enough to provide such a synthesis between **EM** and **GR**. At first glance, his formalism is a conjunction of **GR** and **EM** without mutual interactions between them. The electromagnetic field does not affect the metric in four dimensions, which is a drawback of the theory. His metric does not meet Maudlin's condition (c), i.e. the coupling terms between the unified interactions or an explanation of their mutual effects as a law-like correlation between the two. We shall see that interaction terms do appear in Klein's metric, so for Klein, **EM** does add something to **GR**, and condition (c) is met. Question (III) can't be answered satisfactorily for Kaluza: his theory does not rank high on Maudlin's list as it does not meet condition (c),

[53] Duff (1994, 6).

i.e. it is not a unification à la Maudlin. The mathematical operation that brings in unification does not come with a coupling term.

(B) x^4 as a theoretical entity. A natural argument against Kaluza would be to maintain that the IDs are *ad-hoc* because they are designed to produce the sought after unification. What if the IDs and the dynamics on x^4 are concocted in order to reflect **EM**? As Kaluza extended gravity from 4-D to 5-D, the new dimension seems added to 4-D according to the "letter and spirit" of relativity and thus unification may seem obvious. However, Einstein complained about this extension in a paper he wrote with Grommer in 1923.[54] In **GR**, the covariance of ds^2 was associated with the direct measurability of a 4-D distance. However, there are no similar measures of length or duration in 5-D because "length" here does not have the same meaning. Why should one preserve the covariance of ds^2 in 5-D? Moreover, x^4 is special because *all* fields have the same values along it (which is not the case with $x^0 \ldots x^3$). But there are ways to oppose Einstein's "suspect asymmetry" objection. Firstly, **SR**, a very successful theory, is based on the *asymmetry* between $x^1 \ldots x^3$ and x^0. So there are no logical reasons to refute the fifth dimension on the basis of an asymmetry of the 5-D manifold, as the 4-D manifold is already asymmetric in this sense. Time is already a non-spatial dimension, as the metric along the fourth axis is not a "distance" with a '+' signature.[55] Secondly, the very fact that there is a "measurement operation" associated to four axes does not constitute a necessary condition to be imposed upon other axes. The fifth axis simply acts as a theoretical parameter which has major mathematical significance but which is difficult to measure (physics abounds with such theoretical entities).

(C) The cylinder condition as a brute fact. By imposing CYL, Kaluza tried to accommodate the fact that we do not experience the fifth direction of spacetime.[56] The indication of its existence is only the spectrum of electromagnetic phenomena. Being aware of the outlandish character of this new "extra-world parameter", Kaluza imposed the CYL in order to account for its unobservability.[57] By CYL, the topology of the fifth direction is not affected, it is still the linear topology of $\mathbb{R}$ and the symmetries of spacetime are the same as those of $\mathbf{GR} + \mathbf{EM}$[58] CYL as a brute fact is difficult to tolerate and naturally it seems *ad-hoc*. C. Callender proposes a way to see which facts are brute and which are not: "What we do not want to do is posit substantive truths about the world a priori to meet some unmotivated explanatory demand—as Hegel did when he notoriously said there must be six planets in the solar system."[59] CYL is an unexplained explainer, and a very uncomfortable one.

[54] Einstein and Grommer (1923).
[55] Overduin and Wesson (1998, 3).
[56] This is similar to **SR** which explains why we have the illusion of the "flow of time". Here things seem similar: the universe has a hidden dimension and the illusion or the indication of its existence is the spectrum of electromagnetic phenomena.
[57] "One then has to take into account the fact that we are only aware of the space time variation of quantities, by making their derivatives with respect to the new parameter vanish or by considering them to be small as they are of a higher order." (Kaluza, 1921, 968).
[58] For other details see Duff (1994, 3).
[59] Callender (2004, 206).

For the sake of the beauty and simplicity of his theory, Kaluza committed the same kind of *fiat* that Hegel did.

(D) The scalar field ϕ as a source of "bad" predictions. Novel predictions and observations are respected virtues of a scientific theory and Kaluza simply did not provide any. Actually, Kaluza explicitly refrained from any predictive desideratum when he speculated that his theory did not surpass "mere capricious accident". For macroscopic bodies, Kaluza's theory is a conjunction of **GR** and **EM**. Unlike Weyl's theory, whose unrealistic predictions had scared away Einstein and Pauli, Kaluza's theory did not have blatantly bad predictions on the macroscopic level of observations. Actually, we will see that the new element brought in, ϕ, could possess predictive and explanatory virtues in a specific context and in a specific interpretation. However, for microscopic particles the scalar field gives bad results. For electrons, the interaction with ϕ would be the leading term in (17) and electrons therefore won't follow geodesics. This is blatantly false. Four years before Schrödinger would discover the wave function, Kaluza speculated that, in the future, ϕ could act as a statistical quantity that can explain quantum fluctuations[60] and it *could* get to predictions in the future. Left uninterpreted, ϕ is a troublemaker. But once correctly interpreted, together with its surplus structure, Kaluza's hope was that it would provide explanations to a plethora of phenomena such as the apparent indeterminacy of quantum facts in 4-D. The aim of explaining quantum indeterminism as the appearance of fields existing in extra-dimensions was the Grail of many unified field theories: even Bohr and Einstein coquetted with this idea. But this was mere speculation. What is the scalar ϕ: surplus structure or a would-be *explanans*? For the time being, aside from the approximations of Kaluza's theory, ϕ can be taken as an arbitrary parameter. Without any further interpretation of ϕ, Kaluza's theory seems to be a notational variant of **GR** and **EM** for *macroscopic* objects, acting more as a mathematical formalism than as a physical theory. Although at the beginning of the paper Kaluza exhorted us "[...] to consider *our* space-time to be a four-dimensional part of a $\mathbb{R}^5$",[61] at the end of the paper he became less convincing, downgrading his formalism to a mere computational trickery.[62]

Other than his hope for a future quantum role of ϕ, Kaluza lacks a robust commitment to realism. So one may ask if we do need the fifth dimension more than we need phase space. By analogy, even if phase space is helpful in analytical mechanics, nobody has ever claimed that we really live in a $(q, \dot{q})$ space or (q, p) space. Phase space and configuration space are purely representational spaces that do not produce extra structures such as ϕ. As we have seen, x^4 is similar to such "useful fictions" unless the scalar field ϕ signals the existence of a particle, or it is related to the quantum fluctuations or to the cosmological constant. For Veblen, Hoffmann, as well as for Pais and Jordan, Kaluza's theory was equivalent to a "projective geometry" in which the 4-D manifold was enough and x^4 was projected back to

[60] Kaluza (1921, 865).
[61] Kaluza (1921, 859, my emphasis).
[62] "[...] it is difficult to think that the derived relations, which could scarcely be rejected at the level of theory, represent something more than the enticing game of a capricious chance. If one can establish that the presupposed connections are more than an empty theory, this would be nothing else than a new triumph for Einstein's General Theory of Relativity whose appropriate application to five dimensions has been our concern here." (Kaluza, 1921, 865).

4-D to which they added "vectorial fields" by a "Four-Transformation". One may ask why we do not similarly get rid of the third spatial dimension and just use two dimensions plus a vectorial "height" field. But even if equivalent from a formal point of view to its 3-D counterpart, such a theory will have a hard time in describing and explaining everything in 3D.[63] This will become more perspicuous in the case of Klein's theory. In short, Kaluza's theory illustrates a weak form of unification because it is ad-hoc and does not provide a coupling term between **EM** and **GR**. Moreover, the surplus structure ϕ acts more like a problem maker than a problem solver.

6. EXTRINSIC ELEMENT OF UNIFICATION AND NOVEL EXPLANATIONS IN KLEIN

Klein's new argument and the unification he achieved were more powerful than those of Kaluza's. Klein also employed the IDs as a mathematical procedure, like Kaluza, but he went beyond this. In addressing question (I), I claim that there are two aspects specific to Klein's unification: the extrinsic element of unification and the reduction of types of symmetries of the theory. While the former illustrates the theoretical entity that Morrison demands for unification, the latter is connected to Kitcher's perspective on unification. Both are, I argue, crucial to understanding the improvements upon Kaluza.

(A) The wavefunction as the unification element. Klein's unification element is the behavior of the wavefunction in 5-D which is an extrinsic element to both **GR** and **EM**. It plays the role of the displacement current in Maxwell and it is associated to a mathematical structure, i.e. the Sommerfeld condition of stationarity on a closed orbit. This mathematical condition afterward plays a heuristic role in the discovery of compactification which, as a topological condition, is compatible with both **GR** and **EM**. I want to stress that the wavefunction in 5-D, undoubtedly inspired by de Broglie's Ansatz, is *not* an electromagnetic wave or a gravitational wave *per se*. Being central in the new argument, COMP is a unificatory structure equipped with explanatory powers. It comes from wave mechanics or, from a modern perspective, from the formalism of quantum mechanics in de Broglie's interpretation.

In Klein's case, the unificatory element is part of neither T_1 nor T_2. In trying to answer the second part of question (I), one may ask whether the "extrinsic element of unification" is specific only to Klein's unification. It is worth knowing in general whether the element that generates the theory T_0 is intrinsic to T_1 or to T_2. Klein demonstrates better than any of Morrison's examples the importance of the "extrinsic" element of unification.[64]

(B) Klein's reduction of internal symmetries. Klein was able to explain the symmetries of **EM**, i.e. the internal symmetry expressed by the gauge invariance group $U(1)$,[65]

[63] Thanks to Craig Callender for this suggestion.

[64] Maybe another extrinsic element of unification is the "string" and the "brane" in string theory that are extrinsic elements to both the standard model and to the theory of gravity. I do not claim that an "extrinsic element" of unification characterizes *any* unification.

[65] For a philosophical introduction to "internal symmetries" see Brading and Castellani (2003).

as symmetries of the 5-D manifold. Because of COMP, the symmetry of the **EM** theory is recovered from the symmetry of the spacetime manifold $\mathbb{R}^4 \bigotimes S_1$ and the theory needs only the symmetries of spacetime. We saw that a wave-function invariance demands geometrical transformations associated to the coordinates in 5-D (18). The metric transforms like this:

$$g'_{mn} \rightarrow g_{mn} - \partial_\mu \lambda_\nu - \partial_\nu \xi_\mu \tag{35}$$

and given (21), this corresponds to the gauge invariance symmetry of **EM**:

$$A'^{\mu} \rightarrow A^\mu - \partial_\mu \xi_\mu \tag{36}$$

What is exciting is that $U(1)$ coincides with the invariance on a compactified topology. The *internal* symmetry of **EM** is reduced to $S(1)$ (the symmetry of a one-dimensional manifold) a geometrical consequence of the translation with a multiple of 2π on x^4), which reflects in letter and in spirit the creed of the "geometrization" program. The number of *types* of symmetry is then reduced, and not the sheer *number* of symmetries. This aspect of Klein's theory nicely echoes Kitcher's critique of Friedman's account of unification *qua* explanation.[66] For a bunch of reasons, in the programs inspired by geometrization, spacetime symmetries are preferred to internal symmetries. This reduction/elimination of internal symmetries is manifest in the generalization of Kaluza–Klein to Yang–Mills field and later in string theory: "our spacetime may have extra dimensions and spacetime symmetries in those dimensions are seen as internal (gauge) symmetries from the 4-D point of view. All symmetries could then be unified."[67]

Here gravity and electromagnetism are coupled, unlike Kaluza: in the "line element" (22) there is no longer a pure gravitational "piece of metric". The interaction term $A_\mu A_\nu$ represents the coupling between gravitation and electromagnetism (which Kaluza did not provide) which affects the 4-D gravitational metric. Unlike Kaluza, Klein's theory meets Maudlin's three conditions (a)–(c). Therefore, Klein qualifies as a non-trivial unification which is not a mere conjunction of **GR** and **EM**.

The new manifold is invariant under the group $GL(4) \bigotimes S(1)$, where $S(1)$ is the group of translations, so Klein's theory is not characterizable by a simple group. This is why it does not constitute a high-level unification à la Maudlin. What is more important in this case is the fact that Klein managed to interpret the gauge symmetry of the **EM** theory as a coordinate transformation on the fifth dimension. For Klein the transformations (18) have a dual interpretation: they are coordinate transformations in the full five-dimensional space or as internal gauge transformations in the 4-D spacetime. Some components of the Riemann tensor are interpreted as curvature in 5-D or as field strength in 4-D.[68] Kaluza's theory has the symmetry of a simple group ($GL(5)$), albeit it does not meet condition (c) so

[66] Kitcher (1976).
[67] This is the so called "KK symmetry principle" (Ortín, 2004, 291). Among other meanings, string theorists use unification as reduction of the *types* of symmetries.
[68] O'Raifeartaigh (1997, 50).

gravity and electromagnetism are separated. What Klein managed to do by using COMP is to give a geometrical interpretation to the gauge transformation of **EM**.

(C) Brute facts and explanations. In answering questions (II) and (IV), I claim that the power of explanation in Klein is greatly improved when compared to Kaluza. Klein's reversed argument, in which COMP becomes a brute fact that explains CYL, provided Klein with a powerful unificatory mechanism able to provide *novel* explanations and unexpected predictions. Klein's original intention had been to unify **EM** and **GR** by assuming COMP. The result surpassed his original expectation by explaining the quantization of the electrical charge and the internal symmetry of **EM** as the symmetry of $S(1)$. In addition, there were other unintended, albeit less successful, explanations in Klein's theory. Firstly, it is the mass of the photon. If charge is the component of momentum in the x^4 direction and if one associates a wave to the motion of such a charged particle, the mean value of the field A_μ can be associated to the mass of the photon and $A^{(n)}$ becomes a creation operator after quantization.[69] Although the interpretation of zero modes as masses was too bold for the 1920s, Klein inferred the mass of the photon from $\langle A_\mu \rangle$. For him, as for de Broglie, material particles are solutions to fields and their motion reflects the propagation of waves: "the observed motion as a kind of projection onto space-time of a wave propagation taking place in a space of five dimensions."[70] Klein showed how Schrödinger's equation could be derived from the wave equation in 5-D in which "$\hbar$ does not originally appear, but is introduced in connection with the periodicity in x^4."[71] Does the Planck's constant originate in the periodicity of the fifth dimension? Unfortunately, this is only a partial result, at best. One can infer some quantum numbers, especially the quanta of charge, from the symmetries of x^4, but not *all* of them. How much of quantum theory can be explained by this geometrization program? Not much. Quantum theory in its Hilbert space formulation is not captured by the topology of the fifth dimension,[72] so one should have serious doubts about whether the whole quantum theory can be derived from topological assumptions in extra dimensions. In the eyes of modern physicists, Klein's deduction is flawed: the *classical* theory of fields, even in 5-D, is not able to provide a description of quantum phenomena.

Usually, the major criticism raised against Kaluza–Klein theory is its lack of new predictions. For many physicists, a unification is successful only when making new predictions that are confirmed by experiment.[73] For Kaluza, as well as for **EM** or **GR**, the charge quantization, the symmetry of **EM** and the existence of some particles were brute facts, whereas in Klein's theory they become *explananda*. Once one has accepted COMP, one hits the ground of explanation and no explanation is needed anymore. The "unexplained explainer" is that the fifth dimension is curled and this is for Klein a brute fact such that no other *explanans* is necessary. As part of

[69] van Dongen (2002a, 191).
[70] Klein (1926a, 905).
[71] Klein (1926b).
[72] The question whether a 5-D theory can capture the description of other interpretations of quantum mechanics (Bohmian mechanics, for example) is way beyond the scope of the present paper.
[73] For example, see Smolin (2006, 47, 125) for quotes from Richard Feynman and Sheldon Glashow on superstring theory. Smolin rightly remarks that what used to be critiques against Kaluza–Klein are nowadays directed against string theory.

its unificatory virtues, COMP, a *geometrical* brute fact, explains and predicts *physical* facts. Klein's aim was higher when he envisaged to explain particles. However, can a vacuum theory predict the existence of particles? His theory produced another unexpected explanation: the photon and, albeit Klein was not aware of it, the graviton and the "dilaton" could be deduced from COMP as expectation values of $\langle A_\mu \rangle, \langle g_{\mu\nu} \rangle, \langle \phi \rangle$ by assuming a first-order approximation in which massive states are disregarded, similar to the "dimensional reduction" used in modern Kaluza–Klein theories with $D = 11$ by Scherk, Julia and Cremmer in 1978. The scalar field ϕ, as well as g itself, signals the presence of an *unobservable* particle.

However, the 5-D wavefunction comes with its own troubles: a tower of massive, charged and spin particles with mode $n > 1$ having the mass $m_n = |n|m$ pops into existence. It is easy to see why, in its original formulation, Klein's theory was not renormalizable.[74] Klein's world with a curled x^4 is operationally indistinguishable from a 4-D world with an infinite mass spectrum. The "dimension reduction" is necessary precisely to avoid embarrassing predictions. But in order to explain *massive* particles, one needs non-geometrical fields "coupled" with the metric, which indicates that the geometrical reduction is not fundamental. Despite Klein's attempts, "matter fields" must remain on the brute facts side and cannot be explicated away.

Notwithstanding these shortcomings, when Klein modified Kaluza's original formulation, he was motivated to develop a theory with explanations, with fewer types of brute facts and more capable of solving problems. Klein's case is at odds with Morisson's general decoupling claim: while the wavefunction plays the role of the "theoretical element of unification", Klein's COMP is a mechanism crucial for unification with novel and unexpected explanations, beyond the scope of the original approach (otherwise similar enough to Kaluza). Maxwell had intended to unify electric and magnetic fields, but what he accomplished at the end of the day was the unification of light with electromagnetic waves as well. Similarly, Klein's wavefunction went well beyond its original purpose by providing a mechanism to bridge the theory of classical fields with quantum mechanics.

Last but not least, Klein is a contrast case to Morrison's analysis in still another respect. Morrison tried to show that in Maxwell's unification of **EM**, the theory's commitment to the existence of the ether gradually lessened.[75] As Kaluza–Klein has boosted its explanatory store, the theory illustrates the opposite trend of an increasingly realist commitments to the existence of an extra dimension and to its topology. In Einstein's and Pauli's approaches to extra dimensions, but especially in the later stage of the theory, the realism commitment became more transparent. At the renaissance of the extradimension theory, Cremmer et al. (1981) and Witten (1981) have approached Kaluza–Klein with an explicit realist stance in which the "mechanism" of compactification was based on spontaneous symmetry breaking.[76] From an unexplained explainer, COMP became an *explanandum* of the Kaluza–Klein type of cosmology.

[74] One can associate these massive multiplets with the symmetry group of the theory. According to Salam and Strathdee (1982), the non-compact symmetries are spontaneously broken and they are nothing more than spectrum generating terms.

[75] Morrison (2000, 84)

[76] Appelquist et al. (1987, 278sqq).

(D) Klein's unification: a problem solver and a problem maker. Finally, I want to address question (V) and (VI). Besides the aforementioned explanatory and unificatory boost, COMP acted like a "problem solver" for Kaluza's theory: Klein substantially relaxed the approximation of weak fields, took out the slow motion constraint and showed that electrons move on geodesics. And last but not least, Klein set the ϕ field to a constant, an assumption that eliminates the bad predictions of Kaluza's theory for electrons. This takes his theory to a higher level of unification and in a sense to a different degree of realism.

What are the major limitations of the Kaluza–Klein theory? In the fourth decade of the last century, physicists were preoccupied with the newly discovered nuclear forces. Quantum physics swamped the research in Kaluza–Klein which seemed unable to render a description of these new, quantized interactions.[77] Because of these historical reasons, the Kaluza–Klein program has been stalemated for about half of a century. But notwithstanding these historical reasons and its lack of empirical confirmation, Kaluza–Klein has it own theoretical difficulties.

Aside from the tower of massive particles mentioned above, there are two difficulties generated by COMP. Firstly, the λ_4 is instable. If one "perturbs" the field ϕ, the compactification diverges rapidly either to a singularity or it increases to infinity.[78] Secondly, as directly linked to the elementary charge of the electron, λ_4 is not a dynamic parameter, but a "frozen" parameter of the manifold. This undermines the essence of Einstein's **GR** for whom geometry is dynamical.[79] Klein's theory is dependent then on a background manifold with a fixed topology. Because any particle will have a p^4 momentum dependent on its electrical charge, the Kaluza–Klein model apparently violates the Weak Equivalence Principle, which lies at the foundation of **GR**.[80] Last but not least, from a methodological point of view, the generalization of the Kaluza–Klein theory shows that there are always too many ways to achieve unification. Whenever there are more than one hidden dimensions, there are infinitely many ways to curl them, so there are an infinite number of possible versions of the theory.[81]

7. CONCLUSION

The transition from Kaluza to Klein brought about an increased unificatory and explanatory power, a reduction of types of brute facts while solving previous problems and removing triviality and *ad-hocness*. Although the commitment to

[77] In the meantime, speculations about curled-up extra dimensions seemed "as crazy and unproductive as studying UFOs. There were no implications for experiment, no new predictions, so, in a period when theory developed hand in hand with experiment, no reason to pay attention." (Smolin, 2006, 52). The theory eventually resurfaced due to its generalization to Yang-Mills fields by adding extra dimensions with more and more sophisticated topologies and by including quantum effects.

[78] Smolin (2006, 48).

[79] Smolin (2006, 48).

[80] Wesson (2006, 82).

[81] "The more dimensions, the more degrees of freedom—and the more freedom is accorded to the geometry of the extra dimensions to wander away from the rigid geometry needed to reproduce the forces known in our three-dimensional world." (Smolin, 2006, 51). This inflation of models chases nowadays' string theory, too. The supersymmetric theories are so rich that they can explain almost any imaginable universe. And this affects Kaluza–Klein generalizations which seem to be nothing more than a mathematical tool of representation and not a physical theory that reflects reality.

realism is not transparent in either of these stages, one can see how Klein opened the road to a more realistic interpretation of the higher dimensions. Its potential to be generalized, as well as the paradigmatic mechanism of unification in which internal symmetries are reduced is worthy of further philosophical analysis. The main conclusion of my analysis is that the Kaluza–Klein unification was not possible remaining within the two original theories. The extrinsic factor exploited by Klein, the wavefunction behavior on x^4, is more than a theoretical entity used in unification: it reveals that **GR** and **EM** do not have enough internal resources to be unified. Wave mechanics, or at least a primitive notion of quantum mechanics, was the extrinsic element that endowed the unification with explanatory power.

ACKNOWLEDGEMENTS

I am most grateful to Craig Callender, Nancy Cartwright, Chris Smeenk and Christian Wüthrich for reading various versions of this chapter and for encouraging me to pursue this research. I also want to thank Matt Brown, Tarun Menon, the "Philosophy of Physics Reading Group" at University of California, Irvine, and the audience at the "Second International Conference on the Ontology of Spacetime" (June 9–11, 2006, Montreal, Canada) for some important suggestions. The anonymous referee and the editor for this volume helped with clarifications and corrections on crucial points.

REFERENCES

Aitchison, I.J.R., 1991. The vacuum and unification. In: Saunders, S., Brown, H.J. (Eds.), Philosophy of Vacuum. Clarendon Press, New York, pp. 159–197.

Appelquist, T., Chodos, A., Freund, P.G.O. (Eds.), 1987. Modern Kaluza–Klein Theories. Frontiers in Physics, vol. 65. Addison-Wesley Pub. Co.

Bergmann, P.G., 1942. Introduction to the Theory of Relativity. Prentice-Hall Physics Series, Prentice-Hall.

Brading, K., Castellani, E., 2003. Symmetries in Physics: Philosophical Reflections. Cambridge University Press.

Callender, C., 2004. Measures, explanations and the past: Should 'special' initial conditions be explained? British Journal for the Philosophy of Science 55, 195–217.

Cremmer, E., Julia, B., Scherk, J., 1981. Supergravity theory in eleven-dimensions. Physics Letters B 76, 409.

de Broglie, L., 1924. Thèses présentées a la Faculté des sciences de l'Université de Paris: pour obtenir le grade de docteur ès sciences physiques: soutenues le novembre 1924 devant la Commission d'examen. Masson, Paris.

Ducheyne, S., 2005. Newtons notion and practice of unification. Studies in History and Philosophy of Science 36 (1), 61–78.

Duff, M.J., 1994. Kaluza–Klein theory in perspective. arXiv: hep-th9410046.

Einstein, A., Bergmann, P., 1938. On a generalization of Kaluza's theory of electricity. Annals of Mathematics 39 (3), 683–701.

Einstein, A., Grommer, J., 1923. Beweis der Nichtexistenz eines überall regulären zentrisch symmetrischen Feldes nach der Feld-theorie von Th. Kaluza. Scripta Universitatis atque Bibliothecae Hierosolymitanarum: Mathematica et Physica 1 (1).

Ekspong, G. (Ed.), 1991. The Oskar Klein Memorial Lectures. World Scientific.

Feynman, R.P., Leighton, R.B., Sands, M.L., 1993. The Feynman Lectures on Physics, vol. 2. Mainly Electromagnetism and Matter. Addison-Wesley, Redwood City, CA.
Fock, V., 1926. Über die invariante Form der Wellen- und der Bewegungsgleichungen für einen geladenen Massenpunkt. Zeitschrift für Physik 39 (2), 226–232.
Friedman, M., 1974. Explanation and scientific understanding. Journal of Philosophy 71 (17), 5–19.
Kaku, M., 1994. Hyperspace: A Scientific Odyssey Through Parallel Universes, Time Warps, and the 10th Dimension. Oxford University Press.
Kaluza, T., 1921. In: Zum Unitätproblem der Physik. Sitzungsber. der K. Ak. der Wiss. zu Berlin, pp. 966–972. Translated as "On the unity problem of physics" in Appelquist et al. (1987), pp. 61–69.
Kitcher, P., 1976. Explanation, conjunction, and unification. Journal of Philosophy 73 (22), 207–212.
Kitcher, P., 1981. Explanatory unification. Philosophy of Science 48, 507–531.
Kitcher, P., 1989. In: Explanatory unification and the causal structure of the world. In: Minnesota Studies in the Philosophy of Science, vol. XI. University of Minnesota Press, pp. 410–506.
Klein, O., 1926a. Quantentheorie und fünfdimensionale Relativitätstheorie. Zeitschrift für Physik 37 (12), 895–906. Translated as "Quantum theory and five dimensional theory of relativity" in Appelquist et al. (1987), pp. 76–87.
Klein, O., 1926b. The atomicity of electricity as a quantum theory law. Nature 118, 516.
Klein, O., 1928. Zur fünfdimensionale Darstellung der Relativitätstheorie. Zeitschrift für Physik 46 (3), 188–208.
Lange, M., 2002. An Introduction to the Philosophy of Physics: Locality, Fields, Energy and Mass. Blackwell, Oxford.
Maudlin, T., 1996. On the unification of physics. Journal of Philosophy 93 (3), 129–144.
Morrison, M., 1992. A study in theory unification: The case of Maxwell's electromagnetic theory. Studies in History and Philosophy of Science 23 (1), 103–145.
Morrison, M., 1995. Unified theories and disparate things. Proceedings of the Biennial Meetings of the Philosophy of Science Association 2, 365–373.
Morrison, M., 2000. Unifying Scientific Theories: Physical Concepts and Mathematical Structures. Cambridge University Press.
Norton, J., 1992. Einstein, Nordström and the early demise of scalar, Lorentz-covariant theories of gravitation. Archive for History of Exact Sciences 45 (1), 17–94.
O'Raifeartaigh, L. (Ed.), 1997. The Dawning of Gauge Theory. Princeton University Press.
O'Raifeartaigh, L., Straumann, N., 2000. Gauge theory: Historical origins and some modern developments. Reviews of Modern Physics 72 (1), 1–23.
Ortín, T., 2004. Gravity and Strings. Cambridge University Press.
Overduin, J.M., Wesson, P.S., 1998. Kaluza–Klein gravity. arXiv: gr-qc/0009087.
Plutynski, A., 2005. Explanatory unification and the early synthesis. The British Journal for the Philosophy of Science 56 (3), 595.
Salam, A., Strathdee, J., 1982. On Kaluza–Klein theory. Annals of Physics 141 (2), 316–352.
Schrödinger, E., 1923. Über eine bemerkenswerte Eigenschaft der Quantenbahnen eines einzelnen Elektrons. Zeitschrift für Physik 12 (1), 13–23.
Smolin, L., 2006. The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next. Houghton Mifflin.
Teller, P., 2004. How we dapple the world. Philosophy of Science 71 (4), 425–447.
van Dongen, J., 2002a. Einstein and the Kaluza–Klein particle. Studies in History and Philosophy of Modern Physics 33B (2), 185–210.
van Dongen, J., 2002b. Einstein's unification: General relativity and the quest for mathematical naturalness. Ph.D. thesis, University of Amsterdam.
Vizgin, V., 1994. Unified Field Theories in the First Third of the 20th Century. Birkhäuser, Basel. Translated by Julian B. Barbour (original title: 'Edinye teorii polya v pervoi treti XX veka', Nauka, Moskow, 1985).
Weingard, R., 1984. Grand unified gauge theories and the number of elementary particles. Philosophy of Science 51 (1), 150–155.
Weingard, R., nd. On two goals of unification in physics. Not dated typescript.
Wesson, P.W., 2006. Five-Dimensional Physics: Classical and Quantum Consequences of Kaluza–Klein Cosmology. World Scientific, New Jersey.
Witten, E., 1981. Search for a realistic Kaluza–Klein theory. Nuclear Physics B 186, 412.

CHAPTER 16

Condensed Matter Physics and the Nature of Spacetime

Jonathan Bain*

Abstract

This essay considers the prospects of modeling spacetime as a phenomenon that emerges in the low-energy limit of a quantum liquid. It evaluates three examples of spacetime analogues in condensed matter systems that have appeared in the recent physics literature, indicating the extent to which they are viable, and considers what they suggest about the nature of spacetime.

1. INTRODUCTION

In the philosophy of spacetime literature not much attention has been given to concepts of spacetime arising from condensed matter physics. This essay attempts to address this. It looks at analogies between spacetime and a quantum liquid that have arisen from effective field theoretical approaches to highly correlated many-body quantum systems. Such approaches have suggested to some authors that spacetime can be modeled as a phenomenon that emerges in the low-energy limit of a quantum liquid with its contents (matter and force fields) described by effective field theories (EFTs) of the low-energy excitations of this liquid. In the following, these claims will be evaluated in the context of three examples. Section 2 sets the stage by describing the nature of EFTs in condensed matter systems and how Lorentz-invariance typically arises in low-energy approximations. Section 3 looks at two examples of spacetime analogues in superfluid Helium: analogues of general relativistic spacetimes in superfluid Helium 4 associated with the "acoustic" spacetime programme (e.g., Barceló et al., 2005), and analogues of the Standard Model of particle physics in superfluid Helium 3 (Volovik, 2003). Section 4 looks at a twistor analogue of spacetime in a 4-dimensional quantum Hall liquid (Sparling, 2002). It will be seen that these examples possess limited viability

* Humanities and Social Sciences, Polytechnic University, Brooklyn, USA

The Ontology of Spacetime II
ISSN 1871-1774, DOI: 10.1016/S1871-1774(08)00016-8
© Elsevier BV
All rights reserved

as analogues of spacetime insofar as they fail to reproduce all aspects of the appropriate physics. On the other hand, all three examples may be considered part of a condensed matter approach to quantum gravity; thus to the extent to which philosophers should be interested in concepts of spacetime associated with approaches to quantum gravity, spacetime analogues in condensed matter should be given due consideration.

2. EFFECTIVE FIELD THEORIES IN CONDENSED MATTER SYSTEMS

In general, an effective field theory (EFT, hereafter) of a physical system is a theory of the dynamics of the system at energies close to zero. For some systems, such low-energy states are effectively independent of ("decoupled from") states at high energies. Hence one may study the low-energy sector of the theory without the need for a detailed description of the high-energy sector. Systems that admit EFTs appear in both quantum field theory and condensed matter physics. It is systems of the latter type that will be the focus of this essay.

In particular, the condensed matter systems to be discussed below are highly-correlated quantum many-body systems; that is, many-body systems that display macroscopic quantum effects. Typical examples include superfluids, superconductors, Bose–Einstein condensates, and quantum Hall liquids. The low-energy states described by an EFT of such a system take the form of collective modes of the ground state, generically referred to as "quasiparticles". Such quasiparticles may be either bosonic or fermionic. In the examples below, under the intended interpretation, the latter correspond to the fermionic matter content of spacetime (electrons, neutrinos, etc.), whereas the former correspond to gauge fields (gravitational, electromagnetic, Yang–Mills, etc.) and their quanta (gravitons, photons, etc.).[1] Intuitively, one considers the system in its ground state and tickles it with a small amount of energy. The low-energy ripples that result then take the above forms. To construct an EFT that describes such ripples, the system must first possess an analytically well-defined ground state.[2] An EFT can then be constructed as a *low-energy approximation* of the original theory. One method for doing so is to expand the initial Lagrangian in small fluctuations in the field variables about their ground state values, and then integrate out the high-energy fluctuations. An example of this will be the construction of the EFT for superfluid Helium 4 below.

In this example and the others reviewed in Sections 3 and 4 below, a Lorentz invariant relativistic theory is obtained as the low-energy approximation of a non-relativistic (i.e., Galilei-invariant) theory. Before considering some of the details of

[1] A third type of low-energy state that may arise in condensed matter EFTs takes the form of topological defects of the ground state, the simplest being vortices. This type will not play a role in the following discussion.

[2] In general this is typically not the case. A necessary condition for the existence of an EFT, so characterized, is that the associated system exhibit *gapless* excitations; i.e., low-energy excitations arbitrarily close to the ground state. This notion of an EFT is that described by Polchinski (1993) and Weinberg (1996, p. 145). For Polchinski, an EFT must be "natural" in the sense that all mass terms should be forbidden by symmetries. Mass terms correspond to *gaps* in the energy spectrum insofar as such terms describe excitations with finite rest energies that cannot be made arbitrarily small. For Weinberg, RG theory should only be applied to EFTs that are massless or nearly massless. (Note that this does not entail that massive theories have no EFTs insofar as mass terms that may appear in the high-energy theory may be encoded as interactions between massless effective fields.)

these examples, it may be helpful to get a feel for just how this can come about. It turns out that this is not that uncommon in many non-relativistic condensed matter systems.

Relativistic phenomena are governed by a Lorentz invariant energy dispersion relation of the standard form $E^2 = m^2c^4 + c^2p^2$. This reduces in the massless case to a linear relation between the energy and the momentum: $E^2 = c^2p^2$. It turns out that such a linear relation is a generic feature of the low-energy sector of Bose–Einstein condensates and (bosonic) superfluids. The general form of the dispersion relation for the quasiparticles of these systems is given by $E^2 = c_s^2p^2 + c_s^2p^4/K^2$, where c_s is the quasiparticle speed, and K is proportional to the mass of the constituent bosons (see, e.g., Liberati et al., 2006, p. 3132). In a low-energy approximation, one may assume the quasiparticle momentum is much smaller than the mass of the constituent bosons; i.e., $p \ll K$, and thus obtain a massless relativistic quasiparticle energy spectrum, $E^2 \approx c_s^2p^2$. In Section 3.1, we'll see how this is encoded in the EFT for superfluid Helium 4.

For fermionic quantum liquids, the Fermi surface plays an essential role in the low-energy approximation. For a non-interacting Fermi gas, the Fermi surface is the boundary in momentum space that separates occupied states from unoccupied states and is characterized by the Fermi momentum p_F. In the corresponding EFT, the Fermi surface becomes the surface on which quasiparticle energies vanish. The energy spectrum of low-energy fermionic quasiparticles then goes as $E(p) \approx v_F(p - p_F)$, where $v_F \equiv (\partial E/\partial p)|_{p=pF}$ is the Fermi velocity.[3] This linear dispersion relation suggests the relativistic massless case and figures into the recovery of the relativistic Dirac equation in one- and two-dimensional systems.[4] The massless Weyl equation that describes chiral fermions can also be recovered in the 1-dim case and this will be relevant in the example of a spacetime analogue in a quantum Hall liquid in Section 4. In this example, a $(1 + 1)$-dim relativistic EFT can be constructed for the edge of a 2-dim quantum Hall liquid, and this can then be extended to a $(3 + 1)$-dim EFT for the edge of a 4-dim QH liquid, with an associated notion of spacetime. In three-dimensional systems, the analysis is a bit more complex. An example of a 3-dim system with a relativistic EFT is the A-phase of superfluid Helium 3, which is a fermionic system in which a finite gap exists between the Fermi surface and the lowest energy level, except at two points. When the energy is linearized about these "Fermi points", it takes the form of a dispersion relation formally identical to that for $(3 + 1)$-dim massless relativistic fermions coupled to a 4-potential field that can be interpreted as an electromagnetic potential field. A sketch of the details will be provided in Section 3.2 below.

3. SPACETIME ANALOGUES IN SUPERFLUID HELIUM

This section reviews two examples of spacetime analogues in superfluid Helium: acoustic spacetimes in superfluid Helium 4, and the Standard Model and gravity

[3] Near the Fermi surface, the energy can be linearly expanded as $E(p) = E(p_F) + (\partial E/\partial p)|_{p=p_F}(p - p_F) + \cdots$. Quasiparticle energies vanish on the Fermi surface, hence to second order, $E(p) = v_F(p - p_F)$.

[4] See Zee (2003, p. 274) for the recovery in a system of electrons hopping on a 1-dim lattice, and Zhang (2004, pp. 672–675) for the recovery in current models of 2-dim high temperature superconductors.

in superfluid Helium 3-A. In both of these examples, the low-energy EFT of the system is formally identical to a relativistic theory. The EFT for superfluid Helium 4 is formally identical to a theory describing a massless scalar field in Minkowski spacetime (to first order) or in a curved spacetime (to second order); and the EFT for superfluid Helium 3-A is formally identical to (relevant aspects of) the Standard Model. Associated with these EFTs are concepts of spacetime, and the extent to which the EFTs are adequate analogues of spacetime will depend, in part, on one's prior convictions on how best to model spacetime. For the superfluid Helium examples, these convictions are:

(a) that spacetime is best modeled by (a given aspect of) the solutions to the Einstein equations in general relativity;
(b) that spacetime is best modeled by the ground state for quantum field theories of matter, gauge, and metric fields.

The examples can be judged on the degree to which they reproduce the appropriate physics (general relativity, the Standard Model), as well as the feasibility of the convictions that motivate them. We'll see that spacetime analogues in superfluid Helium 4 are wanting insofar as they do not completely reproduce general relativity, while spacetime analogues in superfluid Helium 3-A are wanting for the same reason, as well as for some qualified reasons concerning the extent to which they reproduce the Standard Model. In the following I will first explain relevant features of each example and then discuss its viability in providing an analogue of spacetime. Section 3.3 will then take up the question of what these examples suggest about the nature of spacetime.

3.1 "Acoustic" spacetimes and superfluid Helium 4

The ground state of superfluid Helium 4 is a Bose–Einstein condensate consisting of ^{4}He atoms (Helium isotopes with four nucleons). It can be characterized by an order parameter that takes the form of a "macroscopic" wavefunction $\varphi_0 = (\rho_0)^{1/2}e^{i\theta}$ with condensate particle density ρ_0 and coherent phase θ. An appropriate Lagrangian describes *non-relativistic* neutral bosons interacting *via* a spontaneous symmetry breaking potential with coupling constant κ (see, e.g., Zee, 2003, pp. 175, 257),

$$\mathcal{L}_{^4\mathrm{He}} = i\varphi^\dagger\partial_t\varphi - \frac{1}{2m}\partial_i\varphi^\dagger\partial_i\varphi + \mu\varphi^\dagger\varphi - \kappa(\varphi^\dagger\varphi)^2, \quad i = 1,2,3. \tag{1}$$

Here m is the mass of a ^{4}He atom, and the term involving the chemical potential μ enforces particle number conservation. This is a thoroughly non-relativistic Lagrangian invariant under Galilean transformations.

A low-energy approximation of (1) can be obtained in a two-step process:[5]

(a) One first writes the field variable φ in terms of density and phase variables, $\varphi = (\rho)^{1/2}e^{i\theta}$, and expands the latter linearly about their ground state values, $\rho = \rho_0 + \delta\rho, \theta = \theta_0 + \delta\theta$ (where $\delta\rho$ and $\delta\theta$ represent fluctuations in density and phase above the ground state).

[5] The following draws on Wen (2004, pp. 82–83) and Zee (2003, pp. 257–258).

(b) After substituting back into (1), one identifies and integrates out the high-energy fluctuations.

Since the ground state φ_0 is a function only of the phase, low-energy excitations take the form of phase fluctuations $\delta\theta$. To remove the high-energy density fluctuations $\delta\rho$, one "integrates" them out: One way to do this is by deriving the Euler–Lagrange equations of motion for the density variable, solving for $\delta\rho$, and then substituting back into the Lagrangian. The result schematically is a sum of two terms: $\mathcal{L}_{^4\mathrm{He}} = \mathcal{L}_0[\rho_0, \theta_0] + \mathcal{L}'_{^4\mathrm{He}}[\delta\theta]$, where the first term describes the ground state of the system and is formally identical to (1), and the second term, dependent only on the phase fluctuations, describes low-energy fluctuations above the ground state. This second term represents the effective field theory of the system and is generally referred to as the effective Lagrangian. To second order in $\delta\theta$, it takes the form,

$$\mathcal{L}'_{^4\mathrm{He}} = \frac{1}{4\kappa}(\partial_t\theta + v_i\partial_i\theta)^2 - \frac{\rho_0}{2m}(\partial_i\theta)^2, \tag{2}$$

with $\delta\theta$ replaced by θ for the sake of notation. Here the second order term depends explicitly on the superfluid velocity $v_i \equiv (1/m)\partial_i\theta$. One now notes that (2) is formally identical to the Lagrangian that describes a massless scalar field in a $(3+1)$-dim curved spacetime:

$$\mathcal{L}'_{^4\mathrm{He}} = \frac{1}{2}\sqrt{-g}\,g^{\mu\nu}\partial_\mu\theta\partial_\nu\theta, \quad \mu,\nu = 0,1,2,3, \tag{3}$$

where the curved effective metric depends explicitly on the superfluid velocity v_i:

$$g_{\mu\nu}\,dx^\mu\,dx^\nu = (\rho/cm)\big[-c^2\,dt^2 + \delta_{ij}\big(dx^i - v^i\,dt\big)\big(dx^j - v^j\,dt\big)\big], \tag{4}$$

where $(-g)^{1/2} \equiv \rho^2/m^2c$, and $c^2 \equiv 2\kappa\rho/m$ (see, e.g., Barceló et al., 2001, pp. 1146–1147). One initial point to note is that, if the original Lagrangian had been expanded to 1st order in $\delta\theta$, the second order term dependent on v_i would vanish in both the effective Lagrangian and the effective metric, and the latter would be formally identical to a flat Minkowski metric (up to conformal constant).[6] This suggests an interpretation of the effective metric (4) as representing low-energy curvature fluctuations (due to the superfluid velocity) above a flat Minkowski background. This is formally identical to the linear approximation of solutions to the Einstein Equations in general relativity, which can likewise be approximated by low-energy fluctuations in curvature above a flat Minkowski background metric. This formal equivalence has been exploited to probe the physics of black holes and the nature of the cosmological constant.

(i) *Acoustic Black Holes*. The general idea is to identify the speed of light in the relativistic case with the speed of low-energy fluctuations, generically referred to as sound modes, in the condensed matter case; hence the terms "acoustic" spacetime and "acoustic" black hole. In general, acoustic black holes are regions in the

[6] When the v_i term is suppressed in (2), the Lagrangian describes a massless field with energy spectrum $E^2 = c^2p^2$. This is the linearly dispersing relation associated with low-energy quasiparticles in Bose–Einstein condensates and bosonic superfluids mentioned in Section 2.

background condensate from which low-energy fluctuations traveling at or less than the speed of sound cannot escape. This can be made more precise with the definitions of acoustic versions of ergosphere, trapped region, and event horizon, among others. A growing body of literature seeks to exploit such formal similarities between relativistic black hole physics and acoustic "dumb" hole physics (see, e.g., Barceló et al., 2005). The primary goal is to provide experimental settings in condensed matter systems for relativistic phenomena such as Hawking radiation associated with black holes.

(ii) *The Cosmological Constant*. Volovik (2003) has argued that the analogy between superfluid Helium and general relativity provides a solution to the cosmological constant problem. The latter he takes as the conflict between the theoretically predicted value of the vacuum energy density in quantum field theory (QFT), and the observational estimate as constrained by general relativity: The QFT theoretical estimate is 120 orders of magnitude greater than what is observed. Volovik sees this as a dilemma for the marriage of QFT with general relativity. If the vacuum energy density contributes to the gravitational field, then the discrepancy between theory and observation must be addressed. If the vacuum energy density is not gravitating, then the discrepancy can be explained away, but at the cost of the equivalence principle. Volovik's preferred solution is to grab both horns by claiming that both QFT and general relativity are EFTs that emerge in the low-energy sector of a quantum liquid.

(a) The first horn is grasped by claiming that QFTs are EFTs of a quantum liquid. As such, the vacuum energy density of the QFT does not represent the true "trans-Planckian" vacuum energy density, which must be calculated from the microscopic theory of the underlying quantum liquid. At $T = 0$, the pressure of such a liquid is equal to the negative of its energy density (Volovik, 2003, pp. 14, 26). This relation between pressure and vacuum energy density also arises in general relativity if the vacuum energy density is identified with the cosmological constant term. However, in the case of liquid ^{4}He in equilibrium, the pressure is zero (Volovik, 2003, p. 29); hence, so is the vacuum energy density.
(b) The second horn is grasped simply by claiming that general relativity is an EFT. Thus, we should not expect the equivalence principle to hold at the "trans-Planckian" level, and hence we should not expect the true vacuum energy density to be gravitating.

Limitations

The implicit claim associated with both the acoustic black hole program and Volovik's solution to the cosmological constant problem is that acoustic spacetimes can be considered analogues of general relativistic spacetimes. One way to assess this claim is by considering the notion of *background structure* in acoustic spacetimes and in general relativity. Note that the acoustic metric arises in a *background-dependent* manner. The acoustic metric (4) is obtained ultimately by imposing particular constraints on *prior* spacetime structure; it is not obtained *ab*

initio.[7] A natural question then is *What should be identified as the background structure of acoustic spacetimes*? The answer to this question will affect the extent to which acoustic spacetimes effectively model general relativity.

One option is to identify Minkowski spacetime as the background structure of acoustic spacetimes. This might be motivated by the explicit form of the acoustic metric (4). As indicated above, it can be interpreted as describing low-energy curvature fluctuations, due to the superfluid velocity, above a flat Minkowski background metric. In particular, (4) can be written in the suggestive form $g_{\mu\nu}\,dx^\mu\,dx^\nu = \eta_{\mu\nu}\,dx^\mu\,dx^\nu + g'_{\mu\nu}\,dx^\mu\,dx^\nu$, where the first term on the right is independent of the superfluid velocity and is identical to a Minkowski metric, and the second term depends explicitly on the superfluid velocity. (The issue of general covariance will be addressed in the subsequent discussion below.)

A second option, however, is to identify the background structure of acoustic spacetimes with (Galilei-invariant) Neo-Newtonian spacetime. This is motivated by considering the procedure by which the acoustic metric was derived. This starts with the Galilei-invariant Lagrangian (1). Low energy fluctuations of the ground state to first order obey the Lorentz symmetries associated with Minkowski spacetime, and low energy fluctuations to second order obey the symmetries of the curved acoustic metric (4).[8] From this point of view, the relation between acoustic spacetimes and Minkowski spacetime is one in which both supervene over a background Neo-Newtonian spacetime. This second option seems the more appropriate: If acoustic metrics are to be interpreted as low-energy fluctuations above the ground state of a condensate, then the background structure of such spacetimes should be identified with the spatiotemporal structure of the condensate ground state, which obeys Galilean symmetries.[9]

This response has implications for the question of the viability of acoustic spacetimes as models of general relativity. Note first that acoustic metrics are not obtained as solutions to the Einstein equations; they are derived *via* a low-energy approximation from the Lagrangian (1) (and similar Lagrangians for other types of condensed matter systems). As noted above, this approximation results schematically in the expansion $\mathcal{L}_{^4\mathrm{He}} = \mathcal{L}_0[\rho_0, \theta_0] + \mathcal{L}'_{^4\mathrm{He}}[\delta\theta]$. To make contact with the Lagrangian formulation of general relativity, Volovik (2003, p. 38) interprets $\mathcal{L}_{^4\mathrm{He}}$ as comprised of a "gravitational" part $\mathcal{L}_0$ describing a background spacetime expressed in terms of the variables θ_0, ρ_0, with gravity being simulated by the superfluid velocity, and a "matter" part $\mathcal{L}'_{^4\mathrm{He}}$, expressed in terms of the variable $\delta\theta$. To obtain the "gravitational" equations of motion, one can proceed in analogy with general relativity by extremizing $\mathcal{L}_{^4\mathrm{He}}$ with respect to θ_0, ρ_0. This results in a set of equations that are quite different in form from the Einstein equations (Volovik, 2003, p. 41), and this indicates explicitly that the *dynamics* of acoustic spacetime EFTs does not reproduce general relativity. Hence acoustic spacetimes cannot be considered *dynamical* analogues of general relativistic spacetimes.

[7] For the moment I will leave aside the question of how this structure can be interpreted. In particular, as will be made explicit in Section 3.3 below, background-dependence of a spacetime theory does not necessarily imply a substantivalist interpretation, any more than background-independence necessarily implies a relationalist interpretation.

[8] Whether or not (4) exhibits non-trivial symmetries will depend on the explicit form of the superfluid velocity.

[9] Again, I will postpone discussion of how this structure can be interpreted until Section 3.3.

While acknowledging that acoustic spacetimes do not model the dynamics of general relativity, some authors have insisted, nonetheless, that acoustic spacetimes account for the *kinematics* of general relativity:

> ... the features of general relativity that one typically captures in an "analogue model" are the *kinematic* features that have to do with how fields (classical or quantum) are defined on curved spacetime, and the *sine qua non* of any analogue model is the existence of some "effective metric" that captures the notion of the curved spacetimes that arise in general relativity. (Barceló et al., 2005, p. 7.)

> The acoustic analogue for black-hole physics accurately reflects half of general relativity—the kinematics due to the fact that general relativity takes place in a Lorentzian spacetime. The aspect of general relativity that does not carry over to the acoustic model is the dynamics—the Einstein equations. Thus the acoustic model provides a very concrete and specific model for separating the kinematic aspects of general relativity from the dynamic aspects. (Visser, 1998, p. 1790.)

Caution should be urged in evaluating claims like these. First, if the kinematics of general relativity is identified with Minkowski spacetime, as linear approximations to solutions to the Einstein equations might suggest, then acoustic spacetimes cannot be considered kinematical analogues of general relativity. And this is because, as argued above, the background structure of acoustic spacetimes should be identified with Neo-Newtonian spacetime and not Minkowski spacetime. More importantly, just what the kinematics of general relativity consists of is open to debate, particularly if we look beyond the linear approximation and consider solutions to the Einstein equations in their full generality. Rather than engage in this debate, I will restrict my comments to two points. First, to the extent that general solutions to the Einstein equations are background independent, they will obviously not be modeled effectively by background dependent acoustic spacetimes. Second, to the extent that the Einstein equations are diffeomorphism invariant, they will not be modeled effectively by acoustic spacetimes, insofar as the low-energy EFT (2) is not diffeomorphism invariant.[10] Thus, insofar as the kinematics of general relativity involves either (or both) of the properties of background independence and diffeomorphism invariance, acoustic spacetimes cannot be said to be kinematical analogues of general relativity.

I would thus submit that acoustic spacetimes provide neither dynamical nor kinematical analogues of general relativity. In fact this sentiment has been expressed in the literature. Barceló et al. (2004) suggest that acoustic spacetimes simply demonstrate that some phenomena typically associated with general relativity really have nothing to *do* with general relativity:

> Some features that one normally thinks of as intrinsically aspects of gravity, both at the classical and semiclassical levels (such as horizons and Hawking radiation), can in the context of acoustic manifolds be instead seen to be

[10] More precisely, the low-energy EFT (2) does not obey "substantive" (as opposed to "formal") general covariance in Earman's (2006, p. 4) sense; i.e., diffeomorphisms are not a local gauge symmetry of (2).

> rather generic features of curved spacetimes and quantum field theory in curved spacetimes, that have nothing to do with gravity *per se*. (Barceló et al., 2004, p. 3.)

This takes some of the initial bite out of Volovik's solution to the cosmological constant problem. If acoustic spacetimes really have nothing to do with general relativity, their relevance to reconciling the latter with QFT is somewhat diminished. While they might provide useful analogues for investigating features of quantum field theory in curved spacetime, extending their use to descriptions of gravitational effects and problems associated with such effects is perhaps not warranted. On the other hand, Volovik's solution to the cosmological constant problem is meant to carry over to other analogues of general relativity besides superfluid ^{4}He. In particular, it can be run for the case of the superfluid ^{3}He-A, which differs significantly from ^{4}He in that fields other than massless scalar fields arise in the low-energy approximation. The fact that these fields model aspects of the dynamics of the Standard Model perhaps adds further plausibility to Volovik's solution. To investigate further, I now turn to ^{3}He.

3.2 The Standard Model and gravity in superfluid Helium 3-A

The second example of a spacetime analogue in a condensed matter system concerns the Standard Model of particle physics and the *A*-phase of superfluid Helium 3. Since ^{3}He atoms are fermions, they can only condense as a Bose–Einstein condensate if they group themselves into bosonic pairs. Thus the particle content of the superfluid consists of pairs of ^{3}He atoms. These pairs are similar to the electron Cooper pairs described by the standard Bardeen–Cooper–Schrieffer (BCS) theory of conventional superconductors. ^{3}He Cooper pairs, however, have additional spin and orbital angular momentum degrees of freedom, and this allows for a number of distinct superfluid phases. In particular, the *A*-phase is characterized by pairs of ^{3}He atoms spinning about anti-parallel axes that are perpendicular to the plane of their orbit.[11]

The (second-quantized) Hamiltonian that describes such ^{3}He-A Cooper pairs takes the following schematic form:

$$H_{^3\text{He-A}} = \chi^\dagger \mathcal{H} \chi, \quad \mathcal{H} = \sigma^b g_b(\vec{p}), \quad b = 1, 2, 3, \tag{5}$$

where the χ's are (non-relativistic) 2-spinors that encode creation and annihilation operators for ^{3}He atoms, σ^a are Pauli matrices, and g_b are three functions of momentum that encode the kinetic energy and interaction potential for ^{3}He-A Cooper pairs.[12] This Hamiltonian can be diagonalized to obtain the quasiparticle

[11] ^{3}He Cooper pairs are characterized by spin triplet ($S = 1$) states with p-wave ($l = 1$) orbital symmetry. There are thus nine distinct types of ^{3}He Cooper pairs, characterized by 3 spin ($S_z = 0, \pm 1$) and 3 orbital ($l_z = 0, \pm 1$) momentum eigenvalues. In ^{3}He-A Cooper pairs, there are no $S_z = 0$ substates, and the orbital momentum axis is aligned with the axis of zero spin.

[12] For details consult Volovik (2003, pp. 82, 96). For inquiring minds, $g_1 = \vec{p} \cdot (\Delta_0 / p_F)(\vec{\sigma} \cdot \hat{d})\hat{m}$, $g_2 = \vec{p} \cdot (\Delta_0 / p_F)(\vec{\sigma} \cdot \hat{d})\hat{n}$, and $g_3 = (p^2/2m) - \mu$. In these expressions, the unit vector d encodes the direction of zero spin, the cross product of the unit vectors m, n encodes the orbital momentum vector, and the constant Δ_0 plays the role of a gap in the BCS energy spectrum for quasiparticle excitations above the Cooper pair condensate. Eq. (5) essentially is a modification of the standard BCS Hamiltonian to account for the extra degrees of freedom of ^{3}He-A Cooper pairs.

energy spectrum. One finds that it vanishes in 3-momentum space at two "Fermi points", call them $p_i^{(a)}$, $i = 1, 2, 3$, $a = 1, 2$. This is due in particular to the directional dependence of the Hamiltonian on the orbital momentum degrees of freedom.

The presence of Fermi points in the energy spectrum is significant for two primarily reasons.[13] First, they are topologically stable insofar as they define singularities in the one-particle Feynman propagator $\mathcal{G} = (ip_0 - \mathcal{H})^{-1}$ that are insensitive to small perturbations. This means pragmatically that the general form of the energy spectrum remains unchanged even when the system undergoes (small) interactions. Second, near the Fermi points $p_\mu^{(a)} = (0, p_i^{(a)})$ in 4-momentum space, the form of the inverse propagator can be expanded as

$$\mathcal{G}^{-1} = \sigma^b e_b^\mu (p_\mu - p_\mu^{(a)}), \quad b = 0, 1, 2, 3 \tag{6}$$

(where the tetrad field e_b^μ encodes the linear approximations of the g_b functions). The quasiparticle energy spectrum is given by the poles in the propagator, and hence takes the general form,

$$g^{\mu\nu}(p_\mu - p_\mu^{(a)})(p_\nu - p_\mu^{(a)}) = 0, \tag{7}$$

where $g^{\mu\nu} = \eta^{ab} e_a^\mu e_b^\nu$. Here the parameters $g^{\mu\nu}$ and $p_\mu^{(a)}$ are dynamical variables insofar as small perturbations of the system are concerned. Again, such perturbations cannot change the fact that Fermi points exist in the energy spectrum; what they *can* change, however, are the positions of the zeros in the energy spectrum, as given by the values of $p_\mu^{(a)}$, or the slope of the curve of the energy spectrum in momentum space, as dictated by the values of $g^{\mu\nu}$.[14]

The Lagrangian corresponding to the energy spectrum (7) can be written as,

$$\mathcal{L}'_{^3\text{He-A}} = \bar{\Psi}\gamma^\mu(\partial_\mu - q^{(a)}A_\mu)\Psi, \tag{8}$$

where $\gamma^\mu = g^{\mu\nu}(\sigma_\nu \otimes \sigma_3)$ are Dirac γ-matrices, the Ψ's are relativistic Dirac 4-spinors (constructed from pairs of the 2-spinors in (5)), and $q^{(a)}A_\mu = p_\mu^{(a)}$. This describes massless Dirac fermions interacting with a 4-vector potential A_μ in a curved spacetime with metric $g_{\mu\nu}$. (8) would be formally identical to the Lagrangian for massless quantum electrodynamics (QED), except for the fact that it does not have a term describing the Maxwell field (i.e., the gauge field associated with the potential A_μ).

It turns out that a Maxwell term arises naturally as a vacuum correction to the coupling between the quasiparticle matter field Ψ and the potential field A_μ. This

[13] The following exposition relies on Volovik (2003, pp. 99–101), and the review in Dreyer (2006, pp. 3–4). Fermi points also occur in the energy spectrum of the sector of the Standard Model above electroweak symmetry breaking (the sector that contains massless chiral fermions). This leads to a theory of universality classes of fermionic vacua based on momentum space topology (Volovik, 2003, Ch. 8). The significance of this theory for the present essay is that superfluid ^{3}He-A and the sector of the Standard Model above electroweak symmetry breaking belong to the same universality class, hence can be expected to exhibit the same low-energy behavior.

[14] This suggests interesting interpretations of the electromagnetic potential and the spacetime metric. To the extent that they can be identified with the objects $p_\mu^{(a)}$ and $g^{\mu\nu}$ in (7), respectively, the electromagnetic potential "... is just the dynamical change in the position of zero in the energy spectrum [of fermionic matter coupled to an electromagnetic field]", and the spacetime metric's role is to change the slope of the energy spectrum (Volovik, 2003, p. 101). The extent to which these identifications are viable is discussed in the following.

is demonstrated by applying the low-energy approximation method outlined in Section 3.1 to the potential field variable: One expands (8) in small fluctuations in A_μ about its ground state value, and then integrates out the high-energy fluctuations. The result is a term that takes the form of the Maxwell Lagrangian in a curved spacetime $\mathcal{L}_{max} = (4\beta)^{-1}\sqrt{-g}\,g^{\mu\nu}g^{\alpha\beta}F_{\mu\alpha}F_{\nu\beta}$, where $F_{\mu\nu}$ is the gauge field associated with the potential A_μ, and β is a constant that depends logarithmically on the cut-off energy.[15] Combining this with (8), the effective Lagrangian for ^{3}He-A then is formally identical to the Lagrangian for massless $(3+1)$-dim QED in a curved spacetime.

Volovik (2003, pp. 114–115) now indicates how this can be extended to include $SU(2)$ gauge fields, and, in principle, the relevant gauge fields of the Standard Model. The trick is to exploit an additional degree of freedom associated with the quasiparticles described by (8). In addition to their charge, such quasiparticles are also characterized by the two values ± 1 of their spin projection onto the axis of zero spin of the underlying ^{3}He-A Cooper pairs. This two-valuedness can be interpreted as a quasiparticle $SU(2)$ isospin symmetry and incorporated explicitly into (8) by coupling Ψ to a new effective field W^i_μ identified as an $SU(2)$ potential field (analogous to the potential for the weak force). Expanding this modified Lagrangian density in small fluctuations in the W-field about the ground state then produces a Yang–Mills term. The general moral is that discrete degeneracies in the Fermi point structure of the energy spectrum induce local symmetries in the low-energy sector of the background liquid (Volovik, 2003, p. 116). For the discrete two-fold ($\mathbb{Z}_2$) symmetry associated with the zero spin axis projection, we obtain a low-energy $SU(2)$ local symmetry; and in principle for larger discrete symmetries $\mathbb{Z}_N$, we should obtain larger local $SU(N)$ symmetry groups. In this way the complete local symmetry structure of the Standard Model could be obtained in the low-energy limit of an appropriate condensed matter system.

Limitations

There are complications to the above procedure, however. The Standard Model has gauge symmetry $SU(3)\otimes SU(2)\otimes U(1)$ with the electroweak sector given by $SU(2)\otimes U(1)$. The electroweak gauge fields belong to non-factorizable representations of $SU(2)\otimes U(1)$, and hence cannot be simply reconstructed from representations of the two separate groups.[16] This suggests that the low-energy EFT of ^{3}He-A does not completely reproduce all aspects of the Standard Model. In fact, it can be demonstrated explicitly that the ^{3}He-A EFT is formally identical only to the sector of the Standard Model above electroweak symmetry breaking, given that both have in common the same Fermi point momentum space topology (see footnote 13).

Moreover, it turns out that general relativity is not fully recovered either. A low-energy treatment of the ^{3}He-A effective metric does not produce the Einstein–Hilbert Lagrangian of general relativity. Under this treatment, one expands the Lagrangian density in small fluctuations in the effective metric $g_{\mu\nu}$

[15] See, e.g., Volovik (2003, p. 112). A detailed derivation is given in Dziarmaga (2002). This method of obtaining the Maxwell term as the second order vacuum correction to the coupling between fermions and a potential field was proposed by Zeldovich (1967).

[16] Thanks to an anonymous referee for making this point explicit.

about the ground state and then integrates out the high-energy terms. This follows the procedure of what is known as "induced gravity", after Sakharov's (1967) derivation of the Einstein–Hilbert Lagrangian density as a vacuum correction to the coupling between quantum matter fields and the spacetime metric. In Sakharov's original derivation, the metric was taken to be Lorentzian, and the result included terms proportional to the cosmological constant and the Einstein–Hilbert Lagrangian density (as well as higher-order terms). In the case of the ^{3}He-A effective metric, the result contains higher-order terms dependent on the superfluid velocity v_i, and these terms dominate the Einstein–Hilbert term.[17] To suppress these terms, Volovik (2003, pp. 130–132) considers the limit in which the mass of the constituent ^{3}He-A atoms goes to infinity (since the superfluid velocity is inversely proportional to this mass, this entails that $v_i \rightarrow 0$). In such an "inert vacuum", the Einstein–Hilbert term can be recovered. Since this limit involves no superflow, Volovik's (2003, p. 113) conclusion is that our physical vacuum cannot be completely modeled by a superfluid. This is suggestive of the formal properties a condensed matter system must possess in order to better model the Standard Model and gravity. In particular, it must possess Fermi points that do not arise *via* symmetry breaking (as the Fermi points of superfluid ^{3}He-A do). From a physical point of view, however, it remains unclear what kind of condensate could possess the property of having infinitely massive constituent particles.

3.3 Interpretation

As has been seen, both of the examples of spacetime analogues in superfluid Helium have their limitations, primarily when it comes to reproducing the relevant physics. To the extent to which they fail to do this, one might question the relevance such examples have to debates over the ontological status of spacetime. On the other hand, both of the above examples, to varying degrees, can be seen to fall within the auspices of a condensed matter approach to quantum gravity. This is explicitly acknowledged by Volovik's (2003) analysis of superfluid ^{3}He-A, and to a lesser extent by the researchers engaged in the acoustic spacetime program (see, e.g., Liberati et al., 2006). This quantum gravity research programme seeks to determine the appropriate condensed matter system that reproduces the matter, gauge and metric fields of current physics in its low-energy approximation, thereby providing a common origin for both quantum field theory and general relativity.[18] Given that all current approaches to quantum gravity are incomplete in one sense or another, the incompleteness of the above examples may thus perhaps be excused. Furthermore, given that philosophers of spacetime should be

[17] See, e.g., Volovik (2003, p. 113). Sakharov's original procedure results in a version of semiclassical quantum gravity, insofar as it describes quantum fields interacting with a classical, unquantized spacetime metric. In the condensed matter context, the background metric is not a classical background spacetime, but rather arises as low-energy degrees of freedom of a quantized non-relativistic system (the superfluid). Hence one could argue this condensed matter version of induced gravity is not semiclassical.

[18] See, e.g., Smolin (2003, pp. 57–58). Thus, to be more precise, the condensed matter programme is an approach to reconciling general relativity and quantum theory, as opposed to an approach to a quantum theory of gravity. Ultimately it suggests gravity need not be quantized, since it claims that gravity emerges in the low energy limit of an already quantized system.

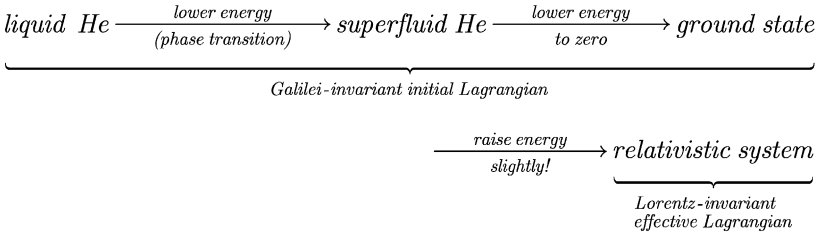

FIGURE 16.1 The relation between the initial Lagrangian and the effective Lagrangian for superfluid Helium.

interested in concepts of spacetime associated with approaches to quantum gravity, they should be interested in concepts of spacetime associated with the above examples, incomplete though they might be. With this attitude in mind, I will now consider what such concepts might look like.

The condensed matter approach to quantum gravity is a background dependent approach to general relativity and the standard model. Under a literal interpretation, it is characterized by the following. First, it suggests that the vacuum of current physics is the Galilei-invariant ground state of a condensate. The Galilei-invariant spatiotemporal structure of the condensate is thus literally interpreted as background spacetime structure. Low-energy collective excitations above the ground state, in the form of fermionic and bosonic quasiparticles, are interpreted as matter, potential, and metric field quanta, respectively; and induced vacuum corrections to the interactions between matter and potential fields are interpreted as gauge fields (the electromagnetic field, the gravitational field, etc.). Note in particular how this picture views violations of Lorentz invariance. It suggests such violations occur at *low* energies, relative to the vacuum; i.e., they occur as one decreases the energy from the realm of the Lorentz-invariant EFT to the Galilei-invariant ground state. Violations of Lorentz-invariance also occur at high energies, relative to the vacuum: they occur as one increases the energy from the realm of the relativistic low-energy EFT to the realm of typical excited states of the condensate. In the example of superfluid Helium, for instance, typical excited superfluid states for temperatures below the critical temperature T_c, will be described by the Galilei-invariant Lagrangian (1). When the energy is increased even more, we eventually pass through the phase transition at T_c and back to the normal liquid state, which, again, is described by the Galilei-invariant Lagrangian (1) (see Figure 16.1).

How this literal interpretation is further qualified; in particular, how one interprets the spatiotemporal structure of the condensate and the nature of, for instance, the low-energy excitation corresponding to the metric field, will depend on one's proclivities, be they relationalist or substantivalist. Let's consider how these further qualifications could play themselves out.

First, any relationalist interpretation should award ontological status *just* to the condensate: relationalists will not countenance interpretations in which the condensate exists in a background spacetime, for instance. A relationalist interpretation might then be based on the following claims:

(1) The background structure consists of the (Galilei-invariant) spatiotemporal relations between the parts of the ground state of the condensate.
(2) Physical fields (matter, gauge, and metric) are low-energy collective excitations of the condensate.
(3) Relativistic spacetime structure consists of the spatiotemporal properties of low-energy excitations.

How Claim (3) gets further qualified may depend on the convictions one possesses on how best to model spacetime in the relativistic context. For instance, if one is convinced that relativistic spacetime is best modeled by the spatiotemporal properties of the ground state for quantum field theories, then one might identify the relativistic spacetime structure of Claim (3) with the spatiotemporal properties of *all* low-energy excitations identified with physical fields. Such convictions underlie a view, common among string theorists, of the relation between general relativity and quantum field theory that prioritizes the latter and Lorentz symmetries. On the other hand, if one is convinced that relativistic spacetime is best modeled by (some aspect of) the solutions to the Einstein equations in general relativity, then the relativistic spacetime structure of Claim (3) might be identified solely with the spatiotemporal properties of that particular low-energy excitation of the condensate identified as the metric field. Convictions of this sort underlie the canonical loop approach to quantum gravity, for which Rovelli (2006) offers a typical relationalist interpretation. Note that having a condensate at the base of everything would make the life of relationalists of the latter stripe a bit more easy. Such relationalists must provide stories that allow them to treat the metric field on par, ontologically, with the other physical fields in nature, and such stories tend to be difficult in the telling (issues such as the non-local nature of the energy associated with the metric field prevent a complete analogy between it and other physical fields, for instance). If there is a condensate substrate common to all physical fields, including the metric field, presumably the latter obtains just as much ontological underpinning from it as the other fields.

Substantivalists of any stripe should award ontological status to both the condensate and spacetime. One can imagine various ways of doing so. A conservative substantivalist, for instance, might adopt the relationalist's Claims (2) and (3) while replacing Claim (1) with

(1′) The background structure consists of the properties of a substantival Neo-Newtonian spacetime.

How Claim (3) gets cashed out by a conservative substantivalist might follow the same maneuvers as the relationalist above. A more intrepid substantivalist might insist on maintaining an ontological distinction between matter and spacetime at all energy scales. One way to do this is to adopt Claims (1′) and (2), but replace (3) with

(3′) Relativistic spacetime structure consists of the properties of a low-energy emergent substantival spacetime.

The full explication of (3′) would require fleshing out a notion of "low-energy emergence". In fact, the examples of low-energy EFTs in superfluid Helium above

(as well as the example in quantum Hall liquids below) have suggested to some authors that novel phenomena including fields, particles, symmetries, twistors, *and* spacetime, *emerge* in the low-energy sector of certain condensed matter systems.[19] Doing justice to this notion of low-energy emergence is perhaps best left to another essay; however, one thing that should be said is that it is distinct from typical notions of emergence associated with phase transitions in condensed matter systems. As Figure 16.1 suggests, typical superfluids can be described by a single Lagrangian that encodes both the normal liquid phase and the superfluid phase, as well as the phase transition between the two. This Lagrangian is formally distinct from the effective Lagrangian of the low-energy sector of the superfluid (when it exists analytically). Thus to the extent that these distinct Lagrangians encode different theories, low-energy emergence can be thought of as a relation between theories, as opposed to a particular interpretation of a single theory (as typical notions of emergence associated with symmetry-breaking phase transitions appear to be).

At this point, it might be appropriate to consider possible motivations for the above substantivalist interpretations. It might not be clear how the roles that typical substantivalists require spacetime to play are accomplished in the condensed matter context. One such role is to provide the ontological substrate for physical fields. Typical substantivalist interpretations of general relativity, for instance, are motivated by a literal interpretation of the representations of physical fields as tensor fields that quantify over the points (or regions) of a differentiable manifold. In the condensed matter context, this intuition might be applied to the field representations of the constituent particles of the condensate as quantifying over the points or regions of Neo-Newtonian spacetime. A conservative substantivalist might claim that, in order to support the condensate, we must postulate the existence of a substantival Neo-Newtonian spacetime. In the case of an intrepid substantivalist, this intuition might be extended to the effective fields of the EFT and the low-energy emergent substantival spacetime; however, it will only do work if the notion of low-energy emergence is cashed out in such a way that the effective fields (and the emergent relativistic spacetime) are sufficiently ontologically distinct from the condensate. Otherwise, relationalists might claim the condensate itself provides the necessary ontological support for the effective fields.

A different type of substantivalist motivation comes from a desire to explain inertial motion in terms of background spacetime structure. A substantivalist might suggest that the coordinated behavior of test particles undergoing inertial motion is mysterious, since such particles have no inertial "antennae" to detect each other, and is explained if we posit a substantival spacetime endowed with an affine con-

[19] In their review of models of analogue gravity, Barceló et al. (2005) speak of "emergent gravitational features in condensed matter systems" (p. 84), and "emergent spacetime symmetries" (p. 89); Dziarmaga (2002, p. 274) describes how "…an effective electrodynamics emerges from an underlying fermionic condensed matter system"; Volovik (2003) in the preface to his text on low-energy properties of superfluid helium, lists "emergent relativistic quantum field theory and gravity" and "emergent non-trivial spacetimes" as topics to be discussed within; Zhang (2004) provides "examples of emergence in condensed matter physics", including the 4-dim quantum Hall effect; and Zhang and Hu (2001, p. 825) speak of the "emergence of relativity" at the edge of 2-dim and 4-dim quantum Hall liquids.

nection that singles out the privileged inertial trajectories.[20] I now want to argue that this motivation for substantivalism fails in the condensed matter context. Note first that it will not do work for a conservative substantivalist. For such a substantivalist, the condensate exists in Neo-Newtonian spacetime, and fields and test particles are low-energy ripples in the condensate. However, the relativistic inertial structure experienced by the ripples is not that possessed by Neo-Newtonian spacetime: According to Claim (3), it consists of the properties of the ripples themselves.[21] An intrepid substantivalist may on first glance fare a bit better: Claim (3′) guarantees that there are substantival privileged inertial trajectories in the relativistic context. In fact, an intrepid substantivalist might even claim to be able to address a key criticism of this motivation; namely, that to explain the origin of inertial motion by referring to privileged inertial trajectories in a substantival spacetime is simply to replace mysterious inertial antennae with mysterious spacetime "feelers" (Brown and Pooley, 2006, p. 72). An intrepid substantivalist might claim to have the basis for an explanation of these feelers: Low-energy ripples of the condensate, viewed as low-energy emergent phenomena, might be expected to coordinate themselves with a low-energy emergent substantival spacetime, given the common origin of the two. Again, whether this basis can be fleshed out into a legitimate explanation will depend on how the notion of low-energy emergence is cashed out. But even if a legitimate explanation in terms of low-energy emergence is forthcoming, it will do no work in distinguishing an intrepid substantivalist from a conservative substantivalist, and hence, in distinguishing substantivalism from relationalism in this context. Note first, that for any notion of low-energy emergence that the intrepid substantivalist adopts, a conservative substantivalist may appropriate it to flesh out Claim (2) and the origin of physical fields. She will then be able to explain the mysterious inertial antennae of such fields in terms of their common substrate origin, to the same degree that the intrepid substantivalist can explain the mysterious spacetime feelers of physical fields in terms of their common origin with (relativistic) spacetime itself. In other words, any legitimate intrepid substantivalist explanation of spacetime feelers will map onto an equally legitimate conservative substantivalist explanation of inertial antennae. And, obviously, a relationalist may engage in the same practice as the conservative substantivalist in this context.

Thus, of the two standard motivations for substantivalism, only the motivation from fields is relevant in the condensed matter context, and intrepid substantivalists will be fairly hard-pressed to make it work for them. Of course this is not to say there may be other motivations for intrepid substantivalism (again, an insistence on a separation between matter and spacetime at all energy scales may be one).[22]

[20] In other words, spacetime has privileged "ruts" along which test particles are constrained to move in the absence of external forces. See, e.g., Brown and Pooley (2006), where this motivation is identified and critiqued.

[21] Note that the "Newtonian limit", $v/c \to 0$, for these relativistic low-energy ripples will consist of non-relativistic low-energy ripples that do experience Neo-Newtonian inertial structure, but again, given the nature of Claim (3), according to a conservative substantivalist, this structure is not to be attributed to the container Neo-Newtonian spacetime, but to properties of the ripples.

[22] For the sake of completeness, two further substantivalist positions can be identified. A super substantivalist might interpret spacetime simply as the condensate itself, with matter fields and gauge fields identified as low-energy aspects of spacetime. Arguably, such a super substantivalist would be hard-pressed to distinguish herself from the relationalist. Both

One might now compare the above notions of spacetime with notions of spacetime associated with other approaches to quantum gravity. Rather than explicitly doing so, the remainder of this section will simply indicate how the condensed matter approach compares conceptually with the two most popular approaches; namely, the background independent canonical loop approach, and background dependent approaches like string theory. The intent is to distinguish these approaches in terms of how they deal with the issues of prior spacetime structure and the nature and status of spacetime symmetries.

(a) The condensed matter approach is distinct from the canonical loop approach, insofar as it is background-*dependent*, the background being the spatiotemporal structure of the condensate. Moreover, while both the condensed matter approach and the loop approach predict violations of Lorentz invariance, these predictions differ in their details. First, as indicated above (see, e.g., Figure 16.1), the condensed matter approach predicts such violations both at low energies (as we approach the ground state), and at high energies (as we approach typical excited states of the condensate and beyond). The loop approach predicts violations only at high energies (scales smaller than the Planck scale) at which it predicts spacetime becomes discrete. Second, the condensed matter approach explains the violation of Lorentz invariance in terms of the existence of a preferred frame; namely the frame defined by the spatiotemporal properties of the condensate, whereas the loop approach explains the violation in terms of background-independence: at the Planck scale, there are no frames, whether Lorentzian or otherwise.[23]

(b) The condensed matter approach differs from background-dependent approaches like string theory in three general respects. First, as is evident in the previous sections, the condensed matter approach differs from string theory in that the structure it attributes to the background is not Minkowskian: Given that the fundamental condensate is a *non-relativistic* quantum liquid, the background will be Neo-Newtonian. Second, while background-dependent approaches that are ultimately motivated by quantum field theory (as string theory is) typically view QFTs as low-energy EFTs of a more fundamental theory, such approaches view the latter as a theory of *high-energy* phenomena (strings, for example). The phenomena of experience, as described by current QFTs, are then interpreted as emerging *via* a process of symmetry breaking. The condensed matter approach, on the other hand, views QFTs and general relativity as EFTs of a more fundamental *low-energy* theory (relative to the vacuum), and the process by which the former arise is a low-energy emergent process that is not to be associated with symmetry breaking. Finally, in general, the condensed matter approach can be characterized by placing less ontological significance on the notion of symmetry than background-dependent approaches in at least two major respects.

First, background-dependent approaches that view QFTs as EFTs describe the phenomena of experience as obeying "imperfect" (gauge) symmetries that result

make the same ontological Claims (1)–(3) and differ only on terminology. A hybrid substantivalist might adopt Claims (1), (2) and (3′); but such a beast would also be hard to motivate: Hybrids cannot consistently appeal to the motivation from fields, given that Claim (1) entails they reject it at the level of the condensate.

[23] Smolin (2003, p. 20) indicates that current experimental data on the violation of Lorentz invariance place very restrictive bounds on preferred frame approaches. Nevertheless he suggests the condensed matter approach may provide key information on the way spacetime might emerge in other scenarios; spin foams, for instance.

from a process of symmetry breaking of a "more perfect" fundamental symmetry. Mathematically, the more perfect fundamental symmetry is hypothesized as having the structure of a single compact Lie group with a minimum of parameters. This is then broken into imperfect symmetries that are characterized by product group structure and relatively many parameters. In particular, the gauge field group structure of the Standard Model, below electroweak symmetry breaking, is given by $U(1)\otimes SU(2)\otimes SU(3)$. In the condensed matter approach, the fundamental condensate is *not* expected to have symmetries described by a single compact Lie group. In the case of superfluid Helium 3, for instance, the "fundamental" symmetries *already* have a "messy" product group structure $U(1)\otimes SO(3)\otimes SO(3)$, reflecting the spin and orbital angular momentum degrees of freedom of ^{3}He Cooper pairs. Moreover, in terms of spacetime symmetries in the condensed matter approach, there are also senses in which the low-energy relativistic (viz., Lorentz) symmetries are *more perfect* than the fundamental Galilean symmetries of the condensate. Note first that the Lorentz group can be characterized as leaving invariant a *single* Lorentzian spacetime metric, whereas the Galilei group cannot; the latter leaves *separate* spatial and temporal metrics invariant. Moreover, the Galilei group does not admit unitary representations, whereas the Lorentz group does.[24]

The second way in which the condensed matter approach de-emphasizes the ontological status of symmetries involves viewing it as an alternative logic of nature to the logic of the Gauge Argument, which typically finds adherents in quantum field theory. According to the Gauge Argument, matter fields are fundamental and imposing local gauge invariance on a matter Lagrangian requires the introduction of interactions with potential gauge fields. The emphasis here is on the fundamental role of local symmetries in explaining the origins of gauge fields (see Martin (2002) for a critique of this argument). According to the condensed matter approach, symmetries, both local and global, as well as matter and potential fields, are low-energy emergent phenomena of the fundamental condensate. In particular, local symmetries do not play a fundamental role in the origin of gauge fields.

4. SPACETIME ANALOGUE IN QUANTUM HALL LIQUIDS

A final example of a spacetime analogue in a condensed matter system concerns the twistor formalism and 4-dimensional quantum Hall liquids. In this example, the low-energy EFT of the edge of the system is formally identical to a theory describing massless relativistic $(3+1)$-dim fields of all helicities. The question of how such a model provides an analogue of spacetime is answered by twistor theory, the goal of which is to reconstruct general relativity and quantum field theory from the conformal properties of twistors. This example is thus similar in spirit with the Helium examples insofar as it, too, can be associated with an approach

[24] Of course these senses depend on a more nuanced characterization of "perfection" in group-theoretic terms than in the case of gauge symmetries. Technically, the second sense is based on the fact that the Galilei group has non-trivial exponents, whereas the Lorentz group does not. Unitary representations of the Galilei group up to a phase factor *can* be constructed (so-called projective representations). The importance of unitary representations comes with implementing spacetime symmetries in the context of quantum theory.

to quantum gravity. Moreover, just as with the Helium examples, this example faces limitations of two types. The first involves the extent to which the model reproduces the "appropriate physics" (which in this case is twistor theory), and the second involves the convictions associated with twistor theory as to how best to represent spacetime (in this case, as derivative of twistors). We'll see that the latter limitation is the most severe: twistor theory faces its own problems in reproducing the appropriate physics (general relativity and quantum field theory). These problems will be discussed in Section 4.4. Section 4.1 describes the context in which 2-dim quantum Hall liquids arise, Section 4.2 indicates how this can be extended to four dimensions, and Section 4.3 explains what twistors are and how they are intended to fit into the picture.

4.1 2-dim quantum Hall liquids

Quantum Hall liquids initially arose in explanations of the 2-dimensional quantum Hall effect (QHE). The set-up consists of current flowing in a 2-dim conductor in the presence of an external magnetic field perpendicular to its surface. The classical Hall effect occurs as the electrons in the current are deflected towards the edge by the magnetic field, thus inducing a transverse voltage. In the steady state, the force due to the magnetic field is balanced by the force due to the induced electric field and the *Hall conductivity* σ_H is given by the ratio of current density to induced electric field. The *quantum* Hall effect occurs in the presence of a strong magnetic field, in which σ_H becomes quantized in units of the ratio of the square of the electron charge e to the Planck constant h:

$$\sigma_H = \nu \times (e^2/h), \tag{9}$$

where ν is a constant. The Integer Quantum Hall Effect (IQHE) is characterized by integer values of ν, and the Fractional Quantum Hall Effect (FQHE) is characterized by values of ν given by odd-denominator fractions. Two properties experimentally characterize the system at such quantized values: The current flowing in the conductor becomes dissipationless, as in a superconductor; and the system becomes incompressible.

These effects can be modeled by a condensate referred to as a quantum Hall (QH) liquid. In one formulation, its constituent particles are represented by "composite" bosons: bosons with p quanta of magnetic flux attached to them, where p is an odd integer.[25] The effect of this coupling is to mimic the Fermi–Dirac statistics of the original electrons. One can show that the total magnetic field felt by the composite bosons vanishes when the constant ν in (9) is given by $1/p$, corresponding to the FQHE. At such values, the bosons feel no net magnetic field, and hence can form a condensate at zero temperature. This condensate, consisting of charged bosons, forms the QH liquid, and can be considered to have the same properties as a superconductor; namely, dissipationless current flow and the expulsion of magnetic fields from its interior. The latter property entails there is no

[25] Technically this is achieved by coupling bosons in the presence of a magnetic field to an additional Chern–Simons field. For details, consult Zhang (1992, p. 32).

net internal magnetic field in a QH liquid, and this entails that the particle density is constant.[26] Thus a QH liquid is incompressible.

The fact that a QH liquid is incompressible entails that there is a finite energy gap between the ground state of the condensate and the first allowable energy states. This means a low-energy approximation cannot be constructed; thus there is no low-energy EFT for the bulk liquid. A low-energy EFT can, however, be constructed for the 1-dim edge of the liquid. Wen (1990) assumed edge excitations take the form of low-energy surface waves and demonstrated that the effective Lagrangian for the edge states describes massless chiral fermion fields in $(1 + 1)$-dim Minkowski spacetime:

$$\mathcal{L}'_{edge} = i\psi^{\dagger}(\partial_t - v\partial_x)\psi, \tag{10}$$

where v is the electron drift velocity.

4.2 4-dim quantum Hall liquids

The $(1+1)$-dim edge Lagrangian (10) tells us little about the ontology of $(3+1)$-dim spacetime. However, it suggests that $(3+1)$-dim massless relativistic fields may be obtainable from the edge states of a 4-dim QH liquid, and this is in fact borne out. Zhang and Hu (2001) provided the first extension of the 2-dimensional QHE to 4-dimensions. In rough outline, they replaced the 2-dim quantum Hall liquid with a 4-dim quantum Hall liquid and then demonstrated that the EFT of the 3-dim edge describes massless fields in $(3 + 1)$-dim Minkowski spacetime.

In slightly more detail, Zhang and Hu made use of a formulation of the 2-dim QHE in terms of spherical geometry first given by Haldane (1983). Haldane considered an electron gas on the surface of a 2-sphere S^2 with a $U(1)$ Dirac magnetic monopole at its center. The radial monopole field serves as the external magnetic field of the original setup. By taking an appropriate thermodynamic limit, the 2-dim QHE on the 2-plane is recovered.[27] Zhang and Hu's extension to 4-dimensions is based on the geometric fact that a Dirac monopole can be formulated as a $U(1)$ connection on a principle fiber bundle $S^3 \rightarrow S^2$, consisting of base space S^2 and bundle space S^3 with typical fiber $S^1 \cong U(1)$ (see, e.g., Nabor, 1997). This fiber bundle is known as the 1st Hopf bundle and is essentially a way of mapping the 3-sphere onto the 2-sphere by viewing S^3 as a collection of "fibers", all isomorphic to a "typical fiber" S^1, and parameterized by the points of S^2. There is also a 2nd Hopf bundle $S^7 \rightarrow S^4$, consisting of the 4-sphere S^4 as base space, and the 7-sphere S^7 as bundle space with typical fiber $S^3 \cong SU(2)$. The $SU(2)$ connection on this bundle is referred to as a Yang monopole. Zhang and Hu's 4-dim QHE then consists of taking the appropriate thermodynamic limit of an electron gas on the surface of a 4-sphere with an $SU(2)$ Yang monopole at its center.

[26] Technically this is due to the fact that the Chern–Simons field is determined by the particle density.

[27] The thermodynamic limit involves taking $N \rightarrow \infty$, $I \rightarrow \infty$, $R \rightarrow \infty$, while holding I/R^2 constant (Haldane, 1983, p. 606; see also Meng, 2003, p. 9415). Here I labels representations of $U(1)$ (and is associated with the Dirac monopole field strength), N is the number of states, which in the lowest energy level is given by $2I+1$, and R is the radius of the 2-sphere. In the lowest energy level, the ratio I/R^2 is proportional to the density of states $N/4\pi R^2$, which must be held constant to recover an incompressible liquid.

Some authors have imbued the interplay between algebra and geometry in the construction of the 4-dim QHE with ontological significance. These authors note that there are only four normed division algebras: the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$, the quaternions $\mathbb{H}$, and the octonions $\mathbb{O}$.[28] It is then observed that these may be associated with the four Hopf bundles, $S^1 \to S^1$, $S^3 \to S^2$, $S^7 \to S^4$, $S^{15} \to S^8$, insofar as the base spaces of these fiber bundles are the compactifications of the respective division algebra spaces $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^4$, $\mathbb{R}^8$. Finally, one notes that the typical fibers of these Hopf bundles are $\mathbb{Z}_2$, $U(1) \cong S^1$, $SU(2) \cong S^3$, and $SO(8) \cong S^7$, respectively. These patterns are then linked with the existence of QH liquids:

> One, two, and four dimensional spaces have the unique mathematical property that boundaries of these spaces are isomorphic to mathematical groups, namely the groups $\mathbb{Z}_2$, $U(1)$ and $SU(2)$. No other spaces have this property. (Zhang and Hu, 2001, p. 827.)

> The four sets of numbers [viz., $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, $\mathbb{O}$] are mathematically known as division algebras. The octonions are the last division algebra, no further generalization being consistent with the laws of mathematics… Strikingly, in physics, some of the division algebras are realized as fundamental structures of the quantum Hall effect. (Bernevig et al., 2003, p. 236803-1.)

> Our work shows that QH liquids work only in certain magic dimensions exactly related to the division algebras… (Zhang, 2004, p. 688.)

These comments have philosophical import to the extent that QH liquids play a fundamental role in physics. They suggest, for instance, an explanation for the dimensionality of space. In particular, if spacetime arises from the edge of a QH liquid, and if the latter only exist in the "magic" dimensions one, two and four, then the spatial dimensions of spacetime are restricted to zero, one, or three, respectively (insofar as the edge would have one less spatial dimension than the bulk). Admittedly, these are big "ifs". The extent to which spacetime arises from the edge of a QH liquid will be dealt with in Sections 4.3 and 4.4 below. The following briefly addresses the extent to which QH liquids can be seen as existing only in a limited number of "magic" dimensions.

Note first that Zhang and Hu's statement should be restricted to the compactifications of the spaces $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^4$, and should include the compactification of $\mathbb{R}^8$ as well, the boundary of the latter being isomorphic to the group $SO(8)$. Furthermore, the statements of Bernevig et al. and Zhang should refer to *normed* division algebras. Baez (2001, p. 149) carefully distinguishes between $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, $\mathbb{O}$ as the only normed division algebras, and division algebras in general, of which there are other examples. Baez (2001, pp. 153–156) indicates how the sequence $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, $\mathbb{O}$ can in principle be extended indefinitely by means of the Cayley–Dickson construction. Starting from an n-dim $*$-algebra A (i.e., an algebra A equipped with a conjugation map $*$), the construction gives a new $2n$-dim $*$-algebra A'. The next member of the sequence after $\mathbb{O}$ is a 16-dim $*$-algebra referred to as the "sedenions". The point here is that the sedenions and all subsequent higher-dimensional

[28] A normed division algebra A is a normed vector space, equipped with multiplication and unit element, such that, for all $a, b \in A$, if $ab = 0$, then $a = 0$ or $b = 0$. $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$ are associative, whereas $\mathbb{O}$ is non-associative (see, e.g., Baez, 2001, p. 149).

constructions do not form division algebras; in particular, they have zero divisors. The question therefore should be whether the absence of zero divisors in a normed $*$-algebra has physical significance when it comes to constructing QH liquids.

Zhang (2004, p. 687) implicitly suggests it does. He identifies various quantum liquids with each Hopf bundle: 1-dim Luttinger liquids[29] with $S^1 \to S^1$, 2-dim QH liquids with $S^3 \to S^2$, and 4-dim QH liquids with $S^7 \to S^4$. Bernevig et al. (2003) complete the pattern by constructing an 8-dim QH liquid as a fermionic gas on S^{15} with an $SO(8)$ monopole at its center. But whether this pattern is physically significant remains to be seen. It is not entirely clear, for example, how the bundle $S^1 \to S^1$, and the trivial $\mathbb{Z}_2$ monopole associated with it, is essential in the construction of Luttinger liquids in general. Moreover, while Luttinger liquids arise at the edge of 2-dim QH liquids, this pattern does not carry over to higher dimensions: it is not the case that 2-dim QH liquids arise at the edge of 4-dim QH liquids, nor is it the case that 4-dim QH liquids arise at the edge of 8-dim QH liquids. Furthermore, and more importantly, Meng (2003) demonstrates that higher-dimensional QH liquids can in principle be constructed for any even dimension, and concludes that the existence of division algebras is not a crucial aspect of such constructions (see also Karabali and Nair, 2002). Hence, while the relation between Hopf bundles and normed division algebras on the one hand, and quantum liquids on the other, is suggestive, it perhaps should not be interpreted too literally.

4.3 Edge states for 4-dim QH liquids and twistors

The low-energy edge states of a 2-dim QH liquid take the form of $(1 + 1)$-dim relativistic massless fields described by (10). These edge excitations can also be viewed as particle-hole dipoles formed by the removal of a fermion from the bulk to outside the QH droplet, leaving behind a hole (see, e.g., Stone, 1990). If the particle-hole separation remains small, such dipoles can be considered single localized bosonic particle states. The stability of such localized states is affected by the uncertainty principle: a stable separation distance entails a corresponding uncertainty in relative momentum, which presumably would disrupt the separation distance. In 1-dim it turns out that the kinetic energy of such dipoles is approximately independent of their relative momentum, hence they are stable. In the case of the 3-dim edge of the 4-dim QH liquid, Zhang and Hu (2001) determined that there is a subset of dipole states for which the isospin degrees of freedom associated with the $SU(2)$ monopole counteract the uncertainty principle. Their main result was to establish that these stable edge states satisfy the $(3 + 1)$-dim zero rest mass field equations for all helicities, and hence can be interpreted as zero rest mass relativistic fields (see also Hu and Zhang, 2002). These include, for instance, spin-1 Maxwell fields and spin-2 graviton fields satisfying the vacuum linearized Einstein equations, as well as massless fields of all higher helicities. Given that there currently is no evidence for the existence of particles with helicities less than

[29] Wen's (1990) EFT (10) identifies the edge of a 2-dim QH liquid as a Luttinger liquid. A Luttinger liquid is comprised of electrons, but differs from a standard Fermi liquid mathematically in the form of the electron propagator. See Wen (2004, pp. 314–315) for details.

2, the latter fact was recognized by Zhang and Hu (2001, p. 827) as an "embarrassment of riches", and a major difficulty of their model.[30]

By itself, this recovery of (3+1)-dim relativistic zero rest mass fields has limited applicability when it comes to questions concerning spacetime ontology. As with the examples in superfluid Helium, we would like to recover general relativity and the Standard Model in their full glory. This is where twistor theory makes its appearance, the goal of which is to recover general relativity and quantum field theory from the structure of zero rest mass fields. Sparling's (2002) insight was to see that Zhang and Hu's stable dipole states correspond to twistor representations of zero rest mass fields. In particular, Sparling demonstrated that the edge of a 4-dim QH liquid can be identified with a particular region of twistor space $\mathbb{T}$. $\mathbb{T}$ is the carrying space for matrix representations of $SU(2,2)$ which is the double covering group of $SO(2,4)$. Elements Z^α of $\mathbb{T}$ are called twistors and are thus spinor representations of $SO(2,4)$. $\mathbb{T}$ contains a Hermitian 2-form $\sum_{\alpha\beta}$ (a "metric") which splits the space into three regions, $\mathbb{T}^+$, $\mathbb{T}^-$, $\mathbb{N}$, consisting of twistors Z^α satisfying $\sum_{\alpha\beta} Z^\alpha Z^\beta > 0$, $\sum_{\alpha\beta} Z^\alpha Z^\beta < 0$, and $\sum_{\alpha\beta} Z^\alpha Z^\beta = 0$, respectively. The connection to spacetime is based on the fact that $SO(2,4)$ is the double covering group of $C(1,3)$, the conformal group of Minkowski spacetime. This allows a correspondence to be constructed under which elements of $\mathbb{N}$, "null" twistors, correspond to null geodesics in Minkowski spacetime, and 1-dim subspaces of $\mathbb{N}$ (i.e., twistor "lines") correspond to Minkowski spacetime points.[31]

To make the identification of the edge of a 4-dim QH liquid with $\mathbb{N}$ plausible, note that the symmetry group of the edge is $SO(4)$ (which is isomorphic to the 3-sphere S^3) and that of the bulk is $SO(5)$ (which is isomorphic to the 4-sphere S^4). The twistor group $SO(2,4)$ contains both $SO(4)$ and $SO(5)$. Intuitively, the restriction of $SO(2,4)$ to $SO(4)$ can be induced by a restriction of twistor space $\mathbb{T}$ to $\mathbb{N}$.[32] With the edge identified as $\mathbb{N}$, edge excitations are identified as deformations of $\mathbb{N}$ (in analogy with Wenn's treatment of the edge in the 2-dim case). In twistor theory, such deformations take the form of elements of the first cohomology group of projective null twistor space $\mathbb{PN}$, and these are in fact solutions to the zero rest mass field equations of all helicities in Minkowski spacetime (Sparling, 2002, p. 25).

Limitations

The complete recovery of twistors from the edge of a 4-dim QH liquid faces a technical hitch concerning the nature of the thermodynamic limit. In the spherical formulations of the QHE, this limit serves to transform the 2-sphere (resp. 4-sphere) into the 2-plane (resp. 4-plane), while reproducing an incompressible QH liquid (footnote 27). In the 4-dim case, this led to Zhang and Hu's "embarrassment of riches" problem: the thermodynamic limit requires taking the isospin

[30] Hu and Zhang (2002, p. 125301-8) consider possible ways to address this problem. A mechanism is needed under which the higher helicity fermionic states acquire masses (i.e., become "gapped") at low energies and thus decouple from observable interactions.

[31] More precisely, the correspondence is between $\mathbb{PN}$, the space of null twistors up to a complex constant (i.e., "projective" null twistors), and *compactified* Minkowski spacetime (i.e., Minkowski spacetime with a null cone at infinity). This is a particular restriction of a general correspondence between projective twistor space $\mathbb{PT}$ and complex compactified Minkowski spacetime. For a brief review, see Bain (2006, pp. 41–42).

[32] Technically, this restriction corresponds to a foliation of the 4-sphere with the level surfaces of the $SO(4)$-invariant function $f(Z^\alpha) = \sum_{\alpha\beta} Z^\alpha Z^\beta$. These surfaces are planes spanned by null twistors (Sparling, 2002, pp. 18–19, 22).

degrees of freedom associated with the Yang monopole to infinity, allowing for $(3+1)$-dim massless fields of all helicities. In the twistor formulation, it is unclear what this limit corresponds to. One way to see this is to note that the twistor formulation does away with the Yang monopole field. In twistor theory, a general result due to Ward allows one to map the dynamics of anti-self-dual Yang–Mills gauge fields (of which the Yang monopole is a particular example) onto purely geometric structures defined on an appropriate twistor space (see, e.g., Bain, 2006, p. 44, for a brief account). Thus in the twistor formulation, there is no explicit isospin space on which to define a limiting procedure. Assumedly, the isospin limit should have a geometrical interpretation in the twistor formulation, but just what it is, is open to speculation (see Sparling, 2002, pp. 27–28 for discussion).

4.4 Interpretation

Even granted that the 4-dim QHE admits a thoroughly twistorial formulation down to the thermodynamic limit, there is still the question of whether spacetime as currently described by general relativity and quantum field theory can be recovered. While Minkowski spacetime can be reconstructed from the space of null twistors, as well as a limited number of field theories, it turns out that no consistent twistor descriptions have been given for massive fields, or for field theories in generally curved spacetimes with matter content. In general, only conformally invariant field theory, and those general relativistic spacetimes that are conformally flat, can be completely recovered in the twistor formalism (see, e.g., Bain, 2006, pp. 45–46 for further discussion). As in the examples of superfluid Helium, one might thus question the relevance that the twistor formulation of the 4-dim QHE has to the ontological status of spacetime. On the other hand, just as with the Helium examples, this twistor example can be viewed as an approach to quantum gravity, and for this reason should be given due consideration.

With this in mind, we may ask what the QH liquid example suggests about the ontological status of spacetime. Taken literally, it suggests that we award fundamental ontological status to a 4-spatial-dimensional quantum Hall liquid. Twistors are then identified as low-energy excitations of the 3-spatial-dimensional edge of this liquid. We then apply the standard practice (and envisioned extensions) of twistor theory to these low-energy excitations to reconstruct spacetime and its contents. On first blush, this interpretation is similar to the superfluid Helium examples in Section 3.3, with twistor theory simply seen as the method for reproducing the relevant physics in the case where the condensed matter system is a QH liquid. Seen in this light, the QH liquid example might be thought to fit within the bounds of Section 3.3's condensed matter approach to quantum gravity. However, the fit is not exact, and consequently how a literal interpretation of the QH liquid example might be further qualified in terms of relationalist and substantivalist options is a bit more nuanced than the superfluid Helium examples. In particular, there are three main differences between the QH liquid example and the superfluid Helium examples.

1. Note first that in the QH liquid example, there is a distinction between the bulk liquid and its edge. Again, spacetime and relativistic field theory are in-

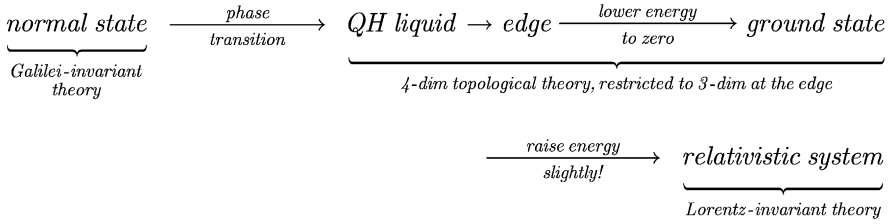

FIGURE 16.2 The relation between theories for a 4-dim quantum Hall liquid.

terpreted as properties, or constructs, of low-energy excitations of the edge (i.e., properties or constructs of twistors), and not of the bulk liquid itself.

2. Second, unlike the superfluid Helium examples, the QH liquid example is not background dependent, at least under one sense of the term. Technically, the theory of a QH liquid is a topological quantum field theory involving a Chern–Simons gauge field.[33] In such a theory, the spacetime metric does not explicitly appear in the term describing the Chern–Simons field (as it does in the Maxwell term in electrodynamics, for instance). Hence the Chern–Simons field does not obey the symmetries of the spacetime metric. Thus, to the extent that background dependence of a theory entails invariance of the theory under the symmetries associated with a particular spatiotemporal structure as encoded in a metric (or set of metrics as in the Galilei case), the theory of a QH liquid is not background dependent. Intuitively, there is no prior *metrical* geometric structure associated with the theory (although there *is* topological/differentiable structure).

3. A third way in which the QH liquid example differs from the superfluid Helium examples concerns the number of theories involved. In the superfluid Helium example, a single theory describes both the normal liquid and the condensate, and this theory is formally distinct from the low-energy EFT (see Figure 16.1). In the QH liquid example, it turns out that the normal state and the condensate are described by different theories, both of which are distinct from the low-energy EFT of the edge (see Figure 16.2). Briefly, the normal liquid is described by a Galilei-invariant theory of electrons moving in a 4-dim conductor, the QH liquid is described by a 4-dim topological theory, and the low-energy EFT of the edge is, in the first instance, a $(3+1)$-dim Lorentz-invariant theory of massless fields of all helicities.[34]

With these qualifications in mind, one can now imagine relationalist and substantivalist interpretations of the QH liquid example. Relationalists should award ontological status just to the QH liquid and may claim:

(1) Physical fields are properties or constructs of low-energy excitations of the edge of the QH liquid.

[33] For the Chern–Simons theory of a 2-dim QH liquid, see Zhang (1992). For the Chern–Simons theory of a 4-dim QH liquid, see Bernevig et al. (2002).

[34] This difference between the two examples is due to the nature of their phase transitions. In the superfluid Helium case, the phase transition is between systems that possess different (internal) symmetries and is characterized by a broken symmetry. In the QH liquid case, the phase transition is between systems that possess different topological orders and is not characterized by a broken symmetry. For a discussion of the notion of topological order, see Wen (2004, Ch. 8).

(2) Relativistic spacetime structure consists of properties or constructs of low-energy excitations of the edge of the QH liquid.

Unlike the superfluid Helium examples, there is no need to further qualify Claim (2), given the convictions of the twistor theorist about the status of relativistic spacetime; i.e., that it's best modeled by twistors, and not by quantum field theory or general relativity.

Substantivalists should award ontological status to both the QH liquid and spacetime. If a substantivalist seeks to ontologically ground the fields that appear in the theory of a QH liquid, she may reify the 4-spatial-dimensional space associated with the liquid. Before the thermodynamic limit is taken, this is a 4-sphere (conceived, not as a metric space, but as a differentiable manifold). Thus a conservative substantivalist might adopt the relationalist's Claims (1) and (2) and add

(3) Spacetime consists of the properties of a substantival differentiable manifold diffeomorphic to the 4-sphere.

An intrepid substantivalist might adopt Claims (1) and (3), qualifying the latter with a restriction to the appropriate energy scale, and replace (2) with

(2′) Relativistic spacetime structure consists of the properties of a low-energy emergent substantival spacetime.

As in the superfluid Helium examples, this would require fleshing out a notion of low-energy emergence. Note that there is still a distinction between the low-energy relativistic EFT of the edge, and the topological theory of the ground state of the edge (see Figure 16.2); hence low-energy emergence might still be considered as a relation between distinct theories. However, the work done by this concept for an intrepid substantivalist in the QH liquid case will be a bit different from the superfluid Helium examples.

Note first that the reasoning in Section 3.3 concerning the typical motivations for substantivalism applies in the QH liquid example as well: The motivation from fields has the potential to do work, whereas that from inertial motion does not. In the case of an intrepid substantivalist in the superfluid Helium examples, the motivation from fields has to be supplemented with an account of low-energy emergence that allows enough of an ontological distinction between emergent fields and spacetime on the one hand, and the underlying condensate on the other to justify the intrepid's claim that (emergent) fields require the existence of (emergent) spacetime for their ontological support. Moreover, low-energy emergence in this context is associated with the low-energy approximation procedure applied directly to the (theory of the) condensate. In the QH liquid example, there is an extra layer of theoretical structure between the condensate and the emergent fields and spacetime; namely, twistors. Thus the emergence associated with spacetime in Claim (2′) above will have to be predicated on the twistor methods that produce spacetime, and at most, only indirectly on the low-energy approximation methods that produce twistors. Thus, again, the intrepid substantivalist has her work cut out for her.

5. CONCLUSION

Interpreting spacetime as a phenomenon that emerges in the low-energy limit of a quantum liquid is problematic for two reasons. First, it depends on the viability of condensed matter analogues of spacetime, and this was seen to be limited in the examples canvassed in this essay. These limitations manifest themselves in a failure to reproduce all aspects of the appropriate physics. For instance, an interpretation of spacetime as emergent in superfluid Helium 4 might be motivated by a desire to model spacetime as (some aspect of) the solutions to the Einstein equations in general relativity. In Section 3.1, we saw that the effective Lagrangian for superfluid Helium 4 lacks both the dynamics associated with general relativity and, arguably, the kinematics. An interpretation of spacetime as emergent in superfluid Helium 3-A might be motivated by a desire to model spacetime as the ground state for quantum field theories of matter, gauge, and metric fields. In Section 3.2, we saw that, while the effective Lagrangian for superfluid ^{3}He-A does reproduce aspects of the Standard Model, it does not reproduce all aspects; nor does it fully recover general relativity. Finally, an interpretation of spacetime as emergent from the edge of a 4-dimensional quantum Hall liquid might be motivated by a desire to derive spacetime using twistor-theoretic techniques. Here the prospects as noted in Section 4.4 are limited primarily by the limitations of twistor theory: Twistor formulations of general solutions to the Einstein equations, and massive interacting quantum fields, have yet to be constructed.

The second way in which interpretations of spacetime as a low-energy emergent phenomenon are problematic has to do with the notion of low-energy emergence itself; in particular, any such interpretation must provide an account of what low-energy emergence is in the condensed matter context. Section 3.3 offered some initial suggestions, however a full account will require significant work. Moreover, we saw in Sections 3.3 and 4.4 that any such notion by itself is compatible with both relationalism and substantivalism. For a relationalist, it would underlie the claim that spatiotemporal structure consists in the spatiotemporal properties of low-energy emergent physical fields; for a substantivalist, it would underlie the claim that spatiotemporal structure consists in the properties of an emergent substantival spacetime. While this latter view might be the most literal way to conceive spacetime as a low-energy emergent phenomenon, arguably it is the hardest to motivate, as Sections 3.3 and 4.4 indicated.

These results suggest that currently an interpretation of spacetime as a low-energy emergent phenomenon cannot be fully justified. However, this essay also argued that such an interpretation should nevertheless still be of interest to philosophers of spacetime. Each of the examples above may be considered part of a general research programme in condensed matter physics; namely, to determine the appropriate condensed matter system that produces the relevant matter, gauge and metric fields of current physics in its low-energy approximation, thus reconciling quantum field theory with general relativity. This research programme may be seen as one path to quantum gravity in competition, for instance, with the background-independent canonical loop approach, and background-dependent approaches like string theory. Thus to the extent that philosophers of spacetime

should consider notions of spacetime associated with approaches to quantum gravity, they should be willing to consider low-energy emergentist interpretations of spacetime.

REFERENCES

Baez, J., 2001. The octonions. Bulletin of the American Mathematical Society 39, 145–205.

Bain, J., 2006. Spacetime structuralism. In: The Ontology of Spacetime, vol. 1. Elsevier, Amsterdam, pp. 37–66.

Barceló, C., Liberati, S., Visser, M., 2001. Analogue gravity from Bose–Einstein condensates. Classical and Quantum Gravity 18, 1137–1156.

Barceló, C., Liberati, S., Visser, M., 2005. Analogue gravity. Living Reviews in Relativity 8 (12). www.livingreviews.org/lrr-2005-12 (cited on 6/9/07).

Barceló, C., Liberati, S., Sonego, S., Visser, M., 2004. Causal structure of acoustic spacetimes. New Journal of Physics 6, 186.

Bernevig, B., Hu, J., Toumbas, N., Zhang, S.-C., 2003. Eight-dimensional quantum Hall effect and "octonions". Physical Review Letters 91, 236803:1–4.

Bernevig, B., Chern, C.-H., Hu, J.-P., Toumbas, N., Zhang, S.-C., 2002. Effective field theory description of the higher dimensional quantum Hall liquid. Annals of Physics 300, 185–207.

Brown, H., Pooley, O., 2006. Minkowski spacetime: A glorious non-entity. In: Dieks, D. (Ed.), The Ontology of Spacetime, vol. 1. Elsevier, Amsterdam, pp. 67–89.

Dreyer, O., 2006. Emergent general relativity. arXiv: gr-qc/0604075.

Dziarmaga, J., 2002. Low-temperature effective electromagnetism in superfluid ^{3}He-A. JETP Letters 75, 273–277.

Earman, E., 2006. The implications of general covariance for the ontology and ideology of spacetime. In: Dieks, D. (Ed.), The Ontology of Spacetime, vol. 1. Elsevier, Amsterdam, pp. 3–24.

Haldane, F.D.M., 1983. Fractional quantization of the Hall effect: A hierarchy of incompressible quantum fluid states. Physical Review Letters 51, 605–608.

Hu, J., Zhang, S.-C., 2002. Collective excitations at the boundary of a four-dimensional quantum Hall droplet. Physical Review B 66, 125301:1–10.

Karabali, D., Nair, V.P., 2002. Quantum Hall effect in higher dimensions. Nuclear Physics B 641, 533–546.

Liberati, S., Visser, M., Weinfurtner, S., 2006. Analogue quantum gravity phenomenology from a two-component Bose–Einstein condensate. Classical and Quantum Gravity 23, 3129–3154.

Martin, C., 2002. Gauge principles, gauge arguments and the logic of nature. Philosophy of Science 69, S221–S234.

Meng, G., 2003. Geometric construction of the quantum Hall effect in all even dimensions. Journal of Physics A 36, 9415–9423.

Nabor, G., 1997. Topology, Geometry, and Gauge Fields. Springer, New York.

Polchinski, J., 1993. Effective field theory and the Fermi surface. arXiv: hep-th/92110046. In: Harvey, J., Polchinski, J. (Eds.), Proceedings of 1992 Theoretical Advanced Studies Institute in Elementary Particle Physics. World Scientific, Singapore, 1993.

Rovelli, C., 2006. The disappearance of space and time. In: Dieks, D. (Ed.), The Ontology of Spacetime, vol. 1. Elsevier, Amsterdam, pp. 25–36.

Sakharov, A.D., 1967. Vacuum quantum fluctuations in curved space and the theory of gravitation. Doklady Akademii Nauk SSSR 177, 70–71. Reprinted in General Relativity and Gravitation 32, 365–367.

Smolin, L., 2003. How far are we from the quantum theory of gravity? arXiv: hep-th/0303185v2.

Sparling, G.A.J., 2002. Twistor theory and the four-dimensional quantum Hall effect of Zhang and Hu. arXiv: cond-mat/0211679.

Stone, M., 1990. Schur functions, chiral bosons, and the quantum Hall effect edge states. Physical Review B 42, 8399–8404.

Visser, M., 1998. Acoustic black holes: Horizons, ergospheres and Hawking radiation. Classical and Quantum Gravity 15, 1767–1791.
Volovik, G., 2003. The Universe in a Helium Droplet. Oxford University Press, Oxford.
Weinberg, S., 1996. The Quantum Theory of Fields, vol. II. Cambridge University Press, Cambridge.
Wen, X.-G., 1990. Chiral Luttinger liquid and the edge excitations in the fractional quantum Hall states. Physical Review B 41, 12838–12844.
Wen, X.-G., 2004. Quantum Field Theory of Many-Body Systems. Oxford University Press, Oxford.
Zee, A., 2003. Quantum Field Theory in a Nutshell. Princeton University Press, Princeton.
Zeldovich, Ya.B., 1967. Interpretation of electrodynamics as a consequence of quantum theory. JETP Letters 6, 345–347.
Zhang, S.-C., 1992. The Chern–Simons–Landau–Ginzberg theory of the fractional quantum Hall effect. International Journal of Modern Physics B 6 (1), 25–58.
Zhang, S.-C., 2004. To see a world in a grain of sand. In: Barrow, J.D., Davies, P.C.W., Harper, C.L. (Eds.), Science and Ultimate Reality: Quantum Theory, Cosmology and Complexity. Cambridge University Press, Cambridge, pp. 667–690.
Zhang, S.-C., Hu, J., 2001. A four-dimensional generalization of the quantum Hall effect. Science 294, 823–828.

SUBJECT INDEX

A
a-boundary, 115–117
Absolute chronological precedence, 62–63
Absolute elements, 136–137, 144
Absolute Foliation, 247, 252
– inscrutability of, 249–250
Absolute Isotropic Expansion, 251, 252
Abstract entities, 42, 44–45
– space points, 42
– time, 44–45
Acceleration, 216–217
Achronal parts, 64–65, 69
– proper, 66
Achronal regions, 62–63, 69, 71
– in Minkowski spacetime, 74–79
Achronal slices, 65, 69, 72, 79
Achronal Universalism, 66–67, 69, 72
Acoustic spacetimes and superfluid Helium 4, 304–309
– acoustic black holes, 305–306
– Cosmological Constant, 306
– limitations, 306–309
Adverbialism, 70
– Minkowskian, 74
Attributes, 3

B
b-boundary, 114–115
b-incompleteness, 113–114
Background fields, 142–145
Background independence, 133–134, 135–138, 144–146, 264, 317
– implications for spacetime ontology, 146–150
Background structures, 134–136
Becoming, 208, 210, 219, 224–226, 229–243, 256
– quantum gravity and, 261
– quantum mechanics and, 236–243
– Special Theory of Relativity and, 229–236, 258–259
– transitivity, 234–236
Being, 258–259, 270
Belnap branching, 189–192
Big Bang, 222, 250–251, 252
Boundaries, 113–117
– *a*-boundary, 115–117
– *b*-boundary, 114–115
Branching spacetime, 187–203
– Belnap branching, 189–192
– ensemble branching, 188–189
– individual branching, 189
– problems with individual branching, 192–203
– – faulty motivation, 192–193
– – implementation difficulties, 193–199
– – shortcomings of the non-Hausdorff option, 199–203
Brans–Dicke theory, 102–104, 107–108
Brute facts, 277–278, 295–297

C
Canonical loop approach, 317
Cauchy surface, 194–195, 201
Causality, 223, 237–238
Change, 43–44, 209, 258
– without time, 266–267
Chern–Simons field, 325
Christoffel symbols, 281–282
Chronogeometric significance, 88–89
Chronology condition, 223
Closed timelike curves (CTCs), 195–196, 197, 225
Compactification of χ^4, 285–291
– charges and matter on geodesics, 287–288
– Klein's metric, 286–287
– the 5-D wavefunction, 288–289
– *See also* Unification
$\text{Compresence}_{\text{EX}}$, 7, 9, 10, 11, 13–14
$\text{Compresence}_{\text{IN}}$, 7, 9, 11
Concurrence, 7
Condensed matter physics, 301, 313–318
– effective field theories in, 302–303
Connection, 91–95, 98, 104
Conservation of energy, 200–201
Conventionality:
– of one-way velocity, 176–178, 179, 183–184
– of simultaneity, 175, 178–180
Cooper pairs, 309, 311
Coordinatization, 141–142

– intrinsic coordinates method, 143
Cosmic Clock Defence, 250, 251
Cosmological Constant, 306
Counterfactual conditionals, 239
Covariance, 139–140, 141–142, 143
Cumulativity, 236, 242
Curve incompleteness, 113
Cylinder condition in Kaluza's unification, 282–283
– as a brute fact, 292

D
Definiteness relation, 230–235
Degeneracy, 207–208, 211, 218, 222, 225–226
Determinism, 154–155, 188, 201
– defining, 156–160
– sophisticated, 161–171
– – arguments against, 165–171
– – correctness of, 163–165
– – rejection of, 171
– violations, 162–163, 165, 168, 171
– *See also* Indeterminism
Diachronic parts, 64–65, 69
– proper, 66
Dicke–Nordtvedt effect, 104
Diffeomorphism invariance, 141–146, 189
Divergence, 189
Doppler effect, 214
Doubly Special Relativity (DSR), 221
Duration, 218, 220, 256–257
– *See also* Time
Dynamical relationalism, 271

E
Effective field theories (EFTs), 301, 304, 306, 317–318
– in condensed matter physics, 302–304
Egalitarianism, 84, 85, 87–88, 99–102, 107
– moderate, 100, 101–102, 107
– strong, 87, 88, 101–102, 107
– weak, 99, 107
Einstein tensor, 98
Electric field, 11–12
Electromagnetic field, 2, 93, 280–281, 282
Electromagnetism, 281
– general relativity, 280–281
– – *See also* Unification
– gravitation and, 280–281
– internal symmetries, 294–295
Elevator thought experiment, 92
Endurance, 60
– definition, 67
– in Galilean spacetime, 69, 70
– in Minkowski spacetime, 72–73
Ensemble branching, 188–189
Entity realism, 19, 20, 22
Epistemic structural realism, 20, 21, 24
– epistemic structural spacetime realism (ESSR), 24
Events, 44–45, 182–183
Exdurance, 60
– definition, 67
– in Galilean spacetime, 69, 70, 71
– in Minkowski spacetime, 72–73, 74
Expanding universe, 250–251
– Absolute Isotropic Expansion, 251, 252
– Law of Uniform Growth, 252
– Relativistic Isotropic Expansion, 251

F
F tropes, 9
Fermi points, 310, 312
Fermi surface, 303
Field kernels, 9–11, 13–14
Field trope-bundle (FTB), 9
– examples of FTB ontology, 11–14
Fields, 5, 8–9, 33
– as properties of a substantial substratum, 1–6
– as trope bundles, 6–11
– background, 142–145
– *See also* Electric field; Electromagnetic field; Gravitational field; Metric field; Scalar field
Fifth dimension, 281–285
– 5-D wavefunction, 288–289, 294, 297
Flat regions, 62, 74–76, 79
Four-dimensional quantum Hall liquids, 320–322
Four-dimensionalism, 60, 177, 258
Fractional Quantum Hall Effect (FQHE), 319
Friedmann–Robertson–Walker (FRW) models, 197

G
G tropes, 9
Galilei-invariant ground state, 313, 325
Gauge Argument, 318
Gauge invariance, 140–141, 146, 147–148, 149
Gauge unification, 279–280
General Theory of Relativity (GTR), 17–18, 27, 84–102, 147, 188–189
– algebraic formulation of general relativity, 126–128
– – generalization, 128
– background independence, 133–134
– branching spacetime and, 193–194, 198–199, 202–203
– dual role of the metric field, 32–35, 47, 49–50, 135

– egalitarian interpretation, 84, 87–88, 99–102, 107
– electromagnetism and, 280–281
– – *See also* Unification
– field interpretation, 84, 85–87, 99, 100, 107
– geometric interpretation, 84, 85–87, 99, 100, 107
– observables, 140–141, 146
– relational nature, 138, 265
– time in, 260
– unification with Quantum Mechanics, 257
Geometrical significance, 88
Geometry, 84
– gravity association, 85–88, 99–102
Global conservation laws, 200
Gravitational field, 47–50, 91–98, 135–136, 146–147, 149, 265
– electromagnetism and, 280–281
– mathematical representatives of, 91–98
– – connection, 91–95
– – metric, 97–98
– – Riemann tensor, 95–97
Gravitational significance, 89
Gravity, 84
– geometry association, 85–88, 99–102
– induced gravity, 312
Growing Block theory, 246

H
Hausdorffness, 119–120, 197, 199
– non-Hausdorff modeling, 199–203
Helium, *See* Superfluid Helium, spacetime analogues
Heraclitean generalization of special relativity, 261–262
Hole argument, 155
– indeterminism, 155, 156–160, 171, 172
Hole diffeomorphism, 155
Hylomorphism, 270

I
Immaterial ether, 2, 3
Immutability, 137
Indefiniteness relation, 231–235
Indeterminism, 155–161, 163, 172, 189–192, 201, 234–235
– branching spacetime and, 192–193
– hole argument, 155, 156–160, 171, 172
– objects playing qualitatively duplicate roles, 168–171
– particle creation, 166–168
– *See also* Determinism
Indexicalism, 70
– Minkowskian, 73–74
Individual branching, 189
Individuation, 270–271
Induced gravity, 312
Inertia, 49–50
Inflationary Big Bang, 250–251, 252
Interger Quantum Hall Effect (IQHE), 319
Intrinsicality, 121–122
Invariance, 139–140, 142–143

K
Kaluza–Klein theory, 275–276
– Kaluza's unification via the 5th dimension, 281–285, 291–294
– – cylinder condition as a brute fact, 292
– – field equations in 5-D, 282
– – geodesics in 5-D, 284–285
– – GR-EM coupling, 291
– – scalar field as a source of bad predictions, 292–294
– – χ^4 as a theoretical entity, 291–292
– Klein's compactification of χ^4, 285–291, 294–298
– – brute facts and explanations, 295–297
– – charges and matter on geodesics, 287–288
– – Klein's metric, 286–287
– – reduction of internal symmetries, 294–295
– – the 5-D wavefunction, 288–289
– – wavefunction as the unification element, 294
– limitations, 298
Kernels, 9–11, 13–14

L
Law of Uniform Growth, 252
Length, 218, 220
– Planck, 221
– proper, 218, 219–220, 226
Local conservation laws, 200
Locality, 119–120, 121–122
Location, 41–43, 60, 63–64
– in Minkowskian spacetime, 72, 74–75
– multilocation, 60
Lorentz cobordance, 196

M
Manifold substantivalism, 22–23, 32–33, 46–49
Mass, 207, 218
Material ether, 3
Mechanical ether, 1–2
Metric, 97–98, 104, 105
Metric field:
– dual role of, 32–35, 47, 49–50, 135
– substantivalism, 22–23, 26–27, 32–33, 49–51
Minimizing the Overall Ontological Revision (MOOR), 78–79

Minkowski spacetime:
– flat and curved achronal regions in, 74–79
– persistence and multilocation in, 71–74
Moments, 44
Motion, 43–44, 146
Multilocation, 60
– in Galilean spacetime, 68–71
– in generic spacetime, 61–68
– in Minkowski spacetime, 71–74

N
No-thin-red-line doctrine, 190–192
Nuclear theory, 9

O
Objective becoming, *See* Becoming
Objects playing qualitatively duplicate roles, 168–171
Observables, 140–141, 146, 148
– complete, 148, 149
– partial, 148, 149
One-way velocity:
– conventionality, 176–180
– impossibility to determine experimentally, 176–177, 180, 184
Ontic structural realism, 20–21, 22, 125
– ontic structural spacetime realism (OSSR), 24–27

P
Parametrized post-Newtonian (PPN) formalism, 102, 106
Particle creation, 166–168
Path, 220
Perdurance, 60
– definition, 67
– in Galilean spacetime, 69, 70, 71
– in Minkowski spacetime, 72–73, 74
Persistence, 59–61
– in Galilean spacetime, 68–71
– in generic spacetime, 61–68
– in Minkowski spacetime, 71–74
Pervasive uniformity, 248
Planck length, 221
Platonia, 263–264
Presentism, 180, 181, 191
Principle of Credulity, 250
Principle of reciprocity, 137
Process, 266–267, 270–271
Proper achronal part, 66
Proper diachronic part, 66
Proper length, 218, 219–220, 226
Proper mass, 207, 218
Proper time, 207, 208, 211, 218–222, 224–226
Property-bundle ontology, 6
Pseudo-Riemannian geometry, 90

Q
Quantum field theory (QTF), 12–14, 202, 306, 317
– time in, 259–260
Quantum gravity, 134, 203, 265
– becoming and, 261
– time in, 259–260
Quantum Hall liquid, 303, 319–320
– 2-D quantum Hall effect (QHE), 319–320
– 4-D quantum Hall liquids, 320–322
– – edge states, 322–323
– spacetime analogue, 318–326, 327
– – interpretation, 324–326
– – limitations, 323–324
Quantum mechanics (QM), 202
– becoming and, 236–243
– relational, 267–269
– time in, 159–160
– unification with General Relativity, 257
Quantum relationalism, 269
Quasiparticles, 302, 303

R
Realism, 19–22
– entity realism, 19, 20, 22
– structural realism, 19–22
– theory realism, 19
– *See also* Structural spacetime realism
Reality:
– four-dimensionalism, 177
– simultaneity and, 180–184
– three-dimensionalism, 180–182
Relational physics, 262–267
– Barbour's Platonia, 263–264
– relational theories of time and space, 262–263
Relational quantum mechanics, 267–269
Relationalism, 23, 51–52, 70, 138–139, 146–147, 264–265
– dynamical, 271
– in the French tradition of epistemology, 269–271
– Minkowskian, 73
– non-reductive, 52–53
– quantum, 269
– reductive, 52
– spacetime relationism, 51–52, 269
– *See also* Substantivalism/relationism debate
Relative mass, 207, 218
Relative time, 207
Relativisation, 270
Relativistic Isotropic Expansion, 251

Relativity:
– Heraclitean generalization of, 261–262
– of simultaneity, 178–179, 180, 181–182, 209–210
– of velocity, 178
– *See also* General Theory of Relativity (GTR); Special Theory of Relativity (STR)
Ricci tensor, 98, 105–106
Riemann tensor, 91, 95–97, 98, 104
Rosen's bimetric theory, 105–108

S
Scalar field, 103–104, 143–144, 292–294
Simultaneity:
– conventionality of, 175, 178–180
– reality and, 180–184
– relativity, 178–179, 180, 181–182, 209–210
Singlet-spin state, 237, 241
Singularities, *See* Spacetime singularities
Sophisticated determinism, 161–171
– arguments against, 165–171
– – objects playing qualitatively duplicate roles, 168–171
– – particle creation, 166–168
– correctness of, 163–165
– rejection of, 171
Sophisticated substantivalism, 53
Space, 256, 264
– dynamization of, 259
– existence of, 40–43
– relational, 262–263
Spacetime, 246, 265
– analogue in quantum Hall liquids, 318–326, 327
– – 2-D quantum Hall liquids, 319–320
– – 4-D quantum Hall liquids, 320–323
– – interpretation, 324–326
– – limitations, 323–324
– analogues in superfluid Helium, 303–318, 327
– – acoustic spacetimes and superfluid Helium 4, 304–309
– – interpretation, 312–318
– – Standard Model and gravity in superfluid Helium 3-A, 309–312
– as a structure, 129–130
– as an ordered set of concrete particulars, 54–55
– existence of, 29, 30–31
– extension, 112–113
– foliation of, 247
– – inscrutability, 249–250
– relationalism, 51–52, 269
– structuralism, 123–125, 129–130, 148
– substantivalism, 45–51
– *See also* Branching spacetime; Structural spacetime realism; Time
Spacetime points, 4, 27, 47
– as abstract entities, 50, 52–56
– as concrete entities, 41, 46, 51, 53–54, 124–125
– definiteness relation, 230–235
– indefiniteness relation, 231–235
Spacetime singularities, 111–112, 197, 198
– algebraic approach, 126–129
– localization, 113–117, 117–118
– – *a*-boundary, 115–117
– – *b*-boundary, 114–115
– non-local aspects, 117–126
– – implications, 120–122
– – structural aspects, 122–126
Spatial part, 69
Spatialized time, 256–258
Special Theory of Relativity (STR), 178, 188, 207
– becoming and, 229–236, 258–259
– Doubly Special Relativity (DSR), 221
– Heraclitean generalization of, 261–262
– passage of time and, 248–249
– three-dimensionalism and, 180–182
– Twin Paradox, 208, 210–216, 219, 224
– – modified, 216–217, 219
Stage theory, 60
Standard Model in superfluid Helium 3-A, 309–312, 318
– limitations, 311–312
String theory, 317
Strong causality principle, 223
Structural realism, 17–18, 19–22, 125
– epistemic, 20, 21, 24
– ontic, 20–21, 22, 125
– *See also* Structural spacetime realism
Structural spacetime realism, 17–18, 23–28
– epistemic (ESSR), 24
– ontic (OSSR), 24–27
– Stein's version of, 28–29
– – objections to, 29–32
Structuralism, 123–125, 129–130, 148–150
Substance, 3, 7, 10
– definition, 25–26, 42
Substance-attribute ontology, 3–6
Substantialism, 270
Substantivalism, 22–28, 44, 46, 135, 146–147, 314–316
– manifold substantivalism, 22–23, 32–33, 46–49
– metric field substantivalism, 22–23, 26–27, 32–33, 49–51
– sophisticated, 53
– spacetime substantivalism, 45–51

Substantivalism/relationism debate, 17–18, 23–28, 32–35, 52, 135–136
– reformulation, 28–32
– singular feature of spacetime and, 122–126
– *See also* Relationalism
Supererogatory actions, 28, 34–35
Superfluid Helium, spacetime analogues, 303–318, 327
– acoustic spacetimes and superfluid Helium 4, 304–309
– – acoustic black holes, 305–306
– – Cosmological Constant, 306
– – limitations, 306–309
– interpretation, 312–318
– Standard Model and gravity in superfluid Helium 3-A, 309–312
– – limitations, 311–312
Symmetry, 317–318
– internal symmetries in electromagnetism, 294–295
– symmetry group, 139–140

T
Temporal becoming, *See* Becoming
Temporal part, 69
Temporality, 256
– *See also* Time
Theory realism, 19
Thin red line, 190–192
Three-dimensionalism, 60, 177, 180–182
Time, 225, 255–256, 259–260
– arguments for unreality, 209–212
– degeneracy, 207–208, 211, 218, 222, 225–226
– existence of, 43–45
– passage of, 246–247
– – Cosmic Clock Defence, 250, 251
– – epistemic objection, 249–251
– – objection from relativistic invariance, 248
– – relativistic invariance, 252
– proper, 207, 208, 211, 218–222, 224–226
– relational, 258, 262–263, 266
– relative, 207
– spatialized, 256–258
– *See also* Spacetime
Time dilation effect, 215
Time lapse, 209, 210–211, 225, 258
Time loops, 223
Time orientability, 194, 195
– dropping, 196–197
Timelessness of physics, 256, 266
Topological cobordance, 196
Topology change, branching spacetime and, 193–199
Tower collapse example, 156–159
Transitivity, 234–236
Trope-bundle ontology, 6–11
– examples, 11–14
Tropes, 6–7
Trouser topology, 63, 193–194, 195
Twin Paradox, 208, 210–216, 219, 224
– modified, 216–217, 219
Twistors, 323, 327
Two-dimensional quantum Hall effect (QFE), 319–320

U
Unicolor, 74–75
Unification, 257, 275–280
– explanatory power and, 277–279
– gauge unification, 279–280
– in theoretical physics, 279–280
– Kaluza's unification via the 5th dimension, 281–285, 291–294
– – cylinder condition as a brute fact, 292
– – field equations in 5-D, 282–284
– – geodesics in 5-D, 284–285
– – GR-EM coupling, 291
– – scalar field as a source of bad predictions, 292–294
– – χ^4 as a theoretical entity, 291–292
– Klein's compactification of χ^4, 285–291, 294–298
– – brute facts and explanations, 295–297
– – charges and matter on geodesics, 287–288
– – Klein's metric, 286–287
– – reduction of internal symmetries, 294–295
– – the 5-D wavefunction, 288–289
– – wavefunction as the unification element, 294
– puzzles of, 276–280
Uniformity, 75

V
Vacuum expectation value (VEV), 12–13
Vacuum hypothesis, 281–282
Velocity:
– conventionality, 176–178, 179, 183–184
– of light, 176–177, 179, 184
– – impossibility to determine experimentally, 176–177, 180, 184
– relativity, 178
Virtuality, 256

W
Weak equivalence principle (WEP), 89
Weyl tensor, 98
Wheeler–De Witt equation, 260–261, 266

X
x trope, 9

Z
Zeno's paradoxes, 43

AUTHOR INDEX

A
Aitchison, I.J.R., 275
Alexander, H.G., 49
Alty, L.J., 195, 197
Amelino-Camelia, G., 221
Anderson, A., 203
Anderson, J.L., 88, 137, 141, 143, 144, 146
Appelquist, T., 297
Aristotle, 43, 180
Arthur, R.T.W., 259

B
Bachelard, G., 269, 270
Baez, J., 321
Bain, J., 127, 323, 324
Balashov, Y., 61, 68, 79
Barbour, J., 225, 262–264, 266
Barceló, C., 301, 306, 308–309, 315
Belnap, N., 189, 190, 192
Belot, G., 5, 32, 34–35, 53, 87, 133, 138, 141, 154, 156, 158, 159–160, 266
Bergmann, P.G., 86, 94, 143, 283, 286
Bergson, H., 256–257, 258–259
Bernevig, B., 321, 325
Bertotti, B., 264
Bigaj, T., 240, 241
Bittner, T., 61, 63
Borde, A., 197
Born, M., 270
Bosshard, B., 115
Brading, K., 294
Brans, C.H., 102–103, 104
Brighouse, C., 153, 154, 155, 156
Broad, C.D., 190
Brown, H., 214, 215, 217, 220, 316
Brown, H.R., 88–89, 91
Budden, T., 96
Bunge, M., 259, 263
Butterfield, J., 49, 66, 118, 122, 153, 154, 155, 156, 261, 266

C
Callender, C., 292
Campbell, K., 7–8
Capek, M., 221, 258–259
Casati, R., 60
Castellani, E., 294
Christian, J., 221, 261
Cleland, C., 122
Clifton, R., 211, 230, 231, 235–236, 237
Connes, A., 267
Costa de Beauregard, O., 258
Crisp, T., 60, 61, 63, 233
Curiel, E., 111, 116, 119, 120–121

D
Daly, C., 7, 10
Damour, T., 104
Davies, P., 210, 213
de Broglie, L., 289
Demaret, J., 118
Descartes, R., 25
DeWitt, B., 203
Dicke, R.H., 102–103, 104, 147
Dieks, D., 30, 134, 208, 259
Dingle, H., 212
DiSalle, R., 18, 30, 35
Donnelly, M., 61, 63
Dorato, M., 18, 21, 30, 121, 124, 210, 214, 235, 237, 238
Douglas, R., 189, 199
Dowker, F., 203
Ducheyne, S., 275
Duff, M.J., 292
Dziarmaga, J., 311, 315

E
Earman, J., 4, 6, 23, 32, 33, 34–35, 39–40, 46, 49, 53, 113, 117, 119, 121, 126, 133, 134, 135, 138, 140, 141, 146, 149–150, 153, 154, 155, 188, 192, 198, 308
Eddington, A.S., 148, 175
Ehlers, J., 89, 93, 98
Einstein, A., 23, 26, 27, 49, 86, 87, 92, 96–98, 136, 137, 140, 147, 149, 175, 176, 177, 216, 220, 280, 283, 290, 292

Ekspong, G., 285, 287, 288
Ellis, G., 114, 195, 223, 225
Eötvös, R.V., 89
Esfeld, M., 17, 20, 21, 25, 26, 124, 125, 242

F
Fauser, B., 259, 260, 261
Faye, J., 40, 44, 54, 231
Feynman, R.P., 88, 89, 100, 101, 276
Field, H., 4, 5
Fierz, M., 100
Finkelstein, J., 240
Flanagan, E.E., 105
Fock, V., 288
Forrest, P., 251
French, S., 17, 20, 21, 124, 125
Friedman, M., 53, 76, 77, 146, 277

G
Gassendi, P., 41
Georgi, H., 280
Geroch, R., 115, 121, 127, 194–195, 197
Ghins, M., 96
Gibbons, G.W., 195
Gibson, I., 61, 63, 71, 74, 75, 76, 79
Gilmore, C., 60, 61, 63, 64, 74
Giulini, D., 94, 144
Gödel, K., 208, 209–211, 219, 222–224, 226
Goenner, H.F.M., 90, 98
Green, G., 1
Greenberger, D., 242
Griffin, D.R., 255–256, 257
Grommer, J., 292
Grünbaum, A., 89, 175, 210, 211

H
Hajicek, P., 200
Haldane, F.D.M., 320
Hall and Hall, 25, 41
Hartle, J.B., 86
Haslanger, S., 60, 70
Hawking, S., 114, 195, 196, 197, 223, 225
Hawley, K., 60, 61
Healey, R., 242
Heathcote, A., 250
Heller, M., 66, 127, 128, 129
Hoefer, C., 23, 32, 33
Hogarth, M., 211, 230, 231, 235–236, 237
Horowitz, G., 121
Howard, D., 242
Hu, J., 315, 320–321, 322–323
Hudson, H., 60, 61, 63, 64

I
Isham, C., 261

J
Jammer, M., 175
Janis, A., 175
Janssen, M., 91–92
Jeans, J., 222
Johnson, R., 115
Jordan, P., 103

K
Kaluza, T., 275–276, 280, 281–285, 291–294
Karabali, D., 322
Kelvin, W.T., 1
Kiefer, C., 259, 260
Kitcher, P., 276, 278, 295
Klein, O., 275–276, 280, 281, 285–291, 294–298
Komar, A., 143
Kretschmann, E., 143
Kuchar, K., 261

L
Ladyman, J., 20, 21, 125
Lam, V., 17, 20, 21, 26, 111, 124, 125
Lange, M., 278
Langton, R., 118
Larmor, J., 2
Lehmkuhl, D., 33
Leibniz, 25, 29, 40, 41
Lewis, D.K., 34, 60, 61, 118, 156, 189, 237–238, 239, 248
Liberati, S., 312
Lorentz, H.A., 2, 3, 218
Lowe, E.J., 42, 55
Lusanna, L., 26, 32

M
MacBride, F., 70
McCabe, G., 199
McCall, S., 189, 190, 192, 193, 198
MacFarlane, J., 190, 192
McGivern, P., 6
Macías, A., 259
Mackie, J., 250
McKinnon, N., 60, 63
Malament, D., 4–5, 175, 179
Martin, C., 318
Mattingly, J., 111
Maudlin, T., 31, 32, 46, 49, 50, 84, 93, 242, 276, 277, 279
Maxwell, N., 210, 230, 278
Melia, J., 153, 155, 156, 157, 158, 160, 161–163, 164–165, 168, 171
Meng, G., 320, 322
Merricks, T., 60, 68
Miller, J.G., 200

Minkowski, H., 177, 218
Misner, C.W., 86, 87–88, 90, 260
Molnar, G., 7, 10, 248
Moreland, J.P., 7
Moriyasu, K., 280
Morrison, M., 275, 278–279, 297
Mukganov, V., 251
Mumford, S.E., 7, 10
Myrvold, W., 231

N
Nabor, G., 320
Nair, V.P., 322
Newton, I., 25, 29, 35, 41
Ni, W.-T., 105
Niiniluoto, I., 19
Nördstrom, G., 280, 282
Nordtvedt, K., 104, 106
Norton, J., 4, 32, 33, 46, 47, 49, 50, 89, 153, 154, 155, 280

O
Ohanian, H., 175
O'Raifeartaigh, L., 279, 286, 287, 295
Ortín, T., 295
Overduin, J.M., 287, 292

P
Parsons, G., 6, 60, 63
Pauli, W., 100
Pauri, M., 18, 26, 32
Peacock, K., 221
Pearle, P., 202
Penrose, R., 189, 199, 203, 263
Petkov, V., 175, 259
Pirani, A.E., 89
Pitts, J.B., 146
Plutynski, A., 275
Poincaré, H., 19–20, 24, 175, 176, 177
Pooley, O., 53–54, 61, 63, 71, 74, 75, 76, 79, 125, 316
Psillos, S., 19
Putnam, H., 210, 211, 230

Q
Quevedo, H., 259

R
Rea, M., 60, 61, 63, 74
Redhead, M., 9, 237
Reichenbach, H., 85, 86, 175, 263
Rendall, A.D., 90
Renn, J., 87, 91–92, 93, 98
Rickles, D., 17, 124, 134, 135, 141, 261
Rietdijk, C.W., 210, 211, 230
Rosen, N., 101, 105–107
Rovelli, C., 27, 33, 52, 85–86, 124, 133, 135, 136, 140, 147, 149, 259, 261, 263, 264–269
Rynasiewicz, R., 5, 18, 23, 136

S
Sachs, R.K., 196
Sakharov, A.D., 312
Salmon, W., 175
Santiago, D.I., 104
Sarkar, S., 175
Sasin, W., 128
Sattig, T., 60, 61, 74
Sauer, T., 87, 93, 98
Saunders, S., 32
Savitt, S., 30, 210, 230, 233, 259
Scheffler, U., 231
Schild, A., 89
Schmidt, B., 114
Schweber, S.S., 13
Scott, S., 115, 116, 117
Senovilla, J., 111
Sider, T., 60, 61, 63, 68, 74
Silbergleit, A.S., 104
Simondon, G., 269, 270–271
Simons, P., 9, 10, 60
Sklar, L., 48, 63, 213, 221
Skow, B., 153, 154, 155, 156, 158, 160, 161–163, 164–165, 167, 168, 171
Slowik, E., 18
Smart, J.J.C., 210, 216
Smith, D., 60, 61, 63
Smith, P., 19
Smolin, L., 137–138, 139, 142, 147, 225, 257, 264–265, 296, 298, 312, 317
Snyder, H.S., 221
Sommerfeld, A., 218–219
Sorkin, R., 197
Souriau, J.-M., 266
Sparling, G.A.J., 301, 323, 324
Spinoza, B., 25
Stachel, J., 17, 32, 125, 149, 175, 271
Stein, H., 18, 25, 28–32, 221–222, 230–234, 248
Stephani, H., 97
Stone, M., 322
Straumann, N., 279, 286, 287
Suárez, M., 21
Swinburne, R., 250
Synge, J.L., 89, 95–96, 148
Szekeres, P., 115, 116, 117

T
Teller, P., 5, 9, 276
Thorne, K.S., 86, 88

Tipler, F.J., 193
Torre, C.G., 141

U
Urchs, M., 231

V
Van Dongen, J., 275, 286, 288, 290, 296
Van Fraassen, B., 19
Varzi, A., 29, 60, 61
Visser, M., 199, 308
Vokrouhlicky, D., 104
Volovik, G., 301, 306, 307, 309, 310, 311, 312, 315

W
Wald, R.M., 86, 90, 97, 100, 101, 113, 179, 193, 194, 198, 200, 260, 265
Wayne, A., 12
Weinberg, S., 103, 278, 302
Weingard, R., 175, 182, 275, 279, 281
Weinstein, S., 103
Wen, X.-G., 304, 320, 322, 325
Wesson, P., 280, 287, 292, 298
Weyl, H., 258, 280, 293
Wheeler, J.A., 86, 88, 138, 140, 142, 147
Wightman, A., 13
Will, C.M., 102, 104, 105, 106
Williams, D.C., 7, 246
Winnie, J., 175
Worrall, J., 20

Y
Yourgrau, P., 211

Z
Zee, A., 303, 304
Zeldovich, Ya.B., 311
Zhang, S.-C., 303, 315, 319, 320–323
Zimmerman, D., 66