WHATEVER IS NEVER AND NOWHERE IS NOT: SPACE, TIME, AND ONTOLOGY IN CLASSICAL AND QUANTUM GRAVITY

by

Gordon Belot

B.Sc., University of Toronto, 1991

M.Sc., University of Toronto, 1993

Submitted to the Graduate Faculty of

University of Pittsburgh in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

1996

John Earman

Joseph Camp
Adolf Grünbaum
John Norton
Carlo Rovelli

WHATEVER IS NEVER AND NOWHERE IS NOT:
SPACE, TIME, AND ONTOLOGY IN CLASSICAL AND QUANTUM GRAVITY

Gordon Belot, Ph.D.

University of Pittsburgh, 1996

Substantivalists claim that spacetime enjoys an existence analogous to that of material bodies, while relationalists seek to reduce spacetime to sets of possible spatiotemporal relations. The resulting debate has been central to the philosophy of space and time since the Scientific Revolution. Recently, many philosophers of physics have turned away from the debate, claiming that it is no longer of any relevance to physics. At the same time, there has been renewed interest in the debate among physicists working on quantum gravity, who claim that the conceptual problems which they face are intimately related to interpretative questions concerning general relativity (GR). My goal is to show that the physicists are correct—there is a close relationship between the interpretative issues of classical and quantum gravity.

In the first part of the dissertation I challenge the received view that substantivalism has a commanding advantage over relationalism on grounds internal to GR. I argue that this view is based on a misconception of the relationships between realism and substantivalism, and between empiricism and relationalism. This has led to a narrow conception of relationalism. Once this is relinquished it can be seen that none of the standard arguments in favor of substantivalism are cogent.

In the second part of the dissertation, I consider the way in which considerations arising out of quantum gravity bear upon the substantival-relational debate. I develop a framework in which to discuss the interpretative problems of gauge theories and place GR in this context. From this perspective, I provide a taxonomy of interpretative options, and

show how the hole argument arises naturally as a consequence of gauge freedom. This means that certain substantivalist interpretations of GR render the theory indeterministic. In the final chapter, I argue that, far from being a drawback, this presents an opportunity for substantivalists. Examples from quantum mechanics, quantum field theory, and quantum gravity, are used to demonstrate that the ambiguities inherent in quantization can lead to an interpretative interplay between theories. In the case of quantum gravity, this means that substantivalism and relationalism suggest, and are suggested by, distinct approaches to quantizing GR.

PREFACE

I suppose that the preface of a dissertation is the appropriate place to discuss the context of discovery and to acknowledge debts acquired along the way. So let me begin by saying that my project had its genesis in a talk which Chris Isham delivered to the Sigma Club in Cambridge during the winter of 1994. At that time, I had been working for several months on the hole argument. This argument, due to John Earman and John Norton, is supposed to show that if one believes that the points of spacetime are real existents, then one is committed to believing that general relativity is an indeterministic theory. It is an ingenious synthesis of Einstein's original hole argument with some themes that have been at the center of philosophy of space and time since the Leibniz-Clarke correspondence. My attitude towards the hole argument had gone through a number stages—which would be, I suspect, quite familiar to anyone who has worked in this area for any length of time. When I was first exposed to it in John Earman's seminar on the philosophy of space and time in the fall of 1991, I had been impressed by its ingenuity—but had been quite sure that it was based on a simple mistake. When I returned to it in the fall of 1993, and began to work through the details of the burgeoning literature of critiques and commentaries focused on the hole argument, I came to appreciate that the situation was considerably more subtle than I had at first thought. This literature contains some very engaging ideas, and I found it quite exciting to pick my way through them in my attempt to reach a stable position, and to understand how the various sorts of response to the hole argument were related to one another and to larger issues within philosophy.

Ultimately, however, I was somewhat dissatisfied with this work. In the introduction to his canonical *World Enough and Spacetime*, Earman had promised that the hole argument would usher in an era of renewed scientific relevance for the philosophy of space and time:

In recent decades the absolute-relational controversy has largely become a captive of academic philosophers. That the controversy is interminably debated in philosophical journals and Ph.D. dissertations is a warning sign that it has lost the relevance to contemporary science that the great natural philosophers of the seventeenth through nineteenth centuries thought it so obviously had for the science of their day. I will attempt to correct this impression by showing, for example, how some of the very concerns raised by Leibniz and Clarke form the core of ongoing foundation problems in the general theory of relativity and how these problems in turn can be used to revitalize what has become an insular and bloodless philosophical discussion. (Earman 1989, p. 3)

Having spent the autumn of 1993 surveying the philosophical literature on the hole argument, I was convinced that Earman's promise was bound to go unfulfilled—this literature consisted of a sizable number of journal articles and a handful of doctoral dissertations, but nowhere was there to be found any substantive connection with ongoing research in physics. Rather, there was a mixture of metaphysical and technical discussion, which, while often ingenious, was just as insular as ever.

It was at this point that I attended Isham's talk "*Prima Facie* Questions in Quantum Gravity" (Isham 1994 is a more technically demanding version of this talk). Isham spoke about the central conceptual difficulty facing attempts to quantize gravity—the fact that time seems to play no role in the quantum theory—and about how this "problem of time" arises in the various approaches to quantum gravity. Along the way he took time to emphasize that he—and many of his colleagues as well—believed that the problem of time was closely related to the hole argument (both following directly from the general covariance of general relativity), and that interpretative questions concerning the ontological status of the spacetime of general relativity had a role to play in determining the best way forward in the face of the formidable technical and conceptual difficulties facing those who were attempting to quantize gravity.

I was very surprised and excited to hear that there were people who thought that the hole argument was relevant to physics after all. I was fortunate to be in the right place at the right time—over the next several months I was able to attend a large number of talks on quantum gravity, at Cambridge and in Durham. I seldom understood very much of what was said. But more than once I was treated to the sight of one or another distinguished physicist leaping up in the middle of a talk in order to denounce the speaker on the grounds that his approach was founded upon the outrageous belief that spacetime points did, or did not, exist.

I decided that I wanted to know how this worked. This dissertation is the result. In the final Chapter I argue that Isham et al are correct—what one believes about the existence of the spacetime points of general relativity influences and is influenced by what one believes about quantum gravity. It takes me a rather long time to reach this conclusion. This is because the way that philosophers think about these issues really has lost touch with the way that physicists think about them. Almost all philosophers of physics seem to believe that it is mandatory to believe that the spacetime points of general relativity enjoy a robust variety of existence. Few physicists would agree. On the other hand, although physicists think of the general covariance of general relativity as being a gauge freedom—and this influences much of what they say about the relation between the interpretative problems of classical and quantum gravity—few philosophers of physics are conversant with this way of talking about things. In order understand the role of interpretative positions in quantum gravity, I had learn to think about general relativity in a new way, and to unlearn some old prejudices. I also, of course, had to learn the rudiments of quantum gravity. This dissertation is a recapitulation of this process—a rather long preamble about the substantival-relational debate and general relativity as a gauge theory, followed by a (relatively) brief discussion of the interplay between interpretative positions in classical and quantum gravity.

This dissertation is not easy to read. The reader is required to change gears a number of times—from a relatively freewheeling survey of the substantival-relational debate to a tightly focused technical discussion; from the technical and conceptual apparatus of general relativity to those of quantum theory, and back again. In particular, the reader is required to bring some knowledge of both the philosophy of quantum mechanics and the philosophy of space and time.

I hope that the reader who has the patience to work through the details will be convinced of two general theses. The first is that substantival-relational debate concerning the nature of the existence of spacetime is, contrary to widespread opinion, far from dead. Indeed, I believe that Earman's promise will be fulfilled—the philosophy of space and time is entering a period of renewed relevance to serious questions of physics. This thesis will be pushed very explicitly throughout the dissertation. My second thesis is that philosophy of science needs to provide an account of contemporary science which makes sense of the multiplicity of subdisciplines of each major discipline, and of the complexity of their interrelationships. This theme will not, for the most part, be made explicit (see, however, §1.4 and the Afterword). Rather, I will be laying out a long and complicated example of the sort of interrelation which exists between contemporary physical theories. But I ask the reader to bear in mind the question of how we must conceive of science if we are to make sense of such examples.

Some of this material is based upon published articles. §§6.2 and 6.3 are derived from "Determinism and Ontology," *International Studies in the Philosophy of Science* 9 (1995): 85-101. The discussion of example 8.4 is based upon "Why General Relativity *Does* Need an Interpretation," forthcoming in PSA 1996. Most of the Appendix is taken directly from "New Work for Counterpart Theorists: Determinism," *British Journal for Philosophy of Science*, 46 (1995): 185-95.

I would like to acknowledge the generous support of the Association of Commonwealth Universities and of the Social Sciences and Humanities Research Council of Canada. I would like to thank Jim Brown and Ian Hacking, who got me started, and the Worm Lady, who kept me going. I would also like to thank my committee in absentia, Jeremy Butterfield and Michael Redhead, under whose direction this project was begun, and without whom it would never have been conceived. I also thank Joe Camp, Adolf Grünbaum, John Norton, and Carlo Rovelli, each of whom, as a member of my committee in Pittsburgh, has offered me guidance, and invaluable comments on drafts and ideas which were often quite chaotic.

Finally, I would like to thank my supervisor, John Earman. Without his good influence on me, and on philosophy of physics in general, this dissertation would never have been possible.

On to the context of justification…

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
The Substantival-Relational Debate in Context

1. The substantival-relational debate is concerned with the nature of the existence of the parts of spacetime. Substantivalists claim that the points of space or spacetime enjoy an existence analogous to that of material substances; relationalists deny that these points are existents, and argue that space or spacetime should be understood as the framework of possible spatiotemporal relations between bodies, rather than as an aggregate of existent points.[1] This debate is perhaps as old as philosophy itself—Aristotle is sometimes identified as a relationalist, arguing in his *Physics* against the substantivalist view of space presented in Plato's *Timaeus* (see, e.g., Sklar 1992, p. 16). But it achieved a new prominence during the Scientific Revolution, when questions about the nature of the existence of space were explicitly tied to questions about the nature of motion. Specifically, it was argued that a substantival space is a precondition for a concept of absolute motion, which is in turn a prerequisite for an empirically adequate dynamical theory (see §2.2). This intermingling of metaphysics and physics gave the substantival-relational debate a unique intellectual flavor, and a sense of relevance to living science.

The debate has remained near the center of philosophical discussion of space and time since the Scientific Revolution.[2] Although there was a brief lull in the decades prior to the inauguration of the so-called 'new philosophy of space and time' by Stein 1967, there has been a great deal of work on the debate in the last thirty years.[3] The publication of

[1] More detailed explications of substantivalism and relationalism will be presented in Chapters 2 and 3. As will become evident, these doctrines are quite difficult to specify. The versions which I will adopt require some substantive discussion by way of motivation. Note that substantivalists and relationalists agree that spacetime itself exists—they disagree only over the nature of the existence of its parts; see §2.3 below.

[2] Friedman 1993 argues that the philosophy of space and time, in turn, has been near the center of philosophical activity since the beginning of the modern period.

[3] Stein 1967 marks a significant departure from earlier work on the philosophy of space and time, both in topic and in approach. It was followed by a series of important works which took up the substantival-relational debate, focusing on substantivalism about spacetime, and framing the discussion as a question of whether or not spacetime deserved to be posited as an unobservable entity, given its explanatory success (see especially Earman

Earman and Norton 1987 marks another watershed. That paper presents a retooled version of Einstein's hole argument, which is supposed to show that if one is a substantivalist about the spacetime of General Relativity (GR) then one is committed to believing that theory to be indeterministic. In recent years there has been a flurry of responses to this argument, setting modern day records for interest in the substantival-relational debate.

This may seem like a portrait of a healthy area of philosophy of physics: here we have a traditional metaphysical problem which has long been recognized to be closely related to problems in the foundations of physics; furthermore, there has recently been a redoubling of philosophical interest in response to an argument which ties the debate to important questions about one of our best physical theories. An examination of the hole argument literature reveals, however, that many philosophers of physics are skeptical about the interest and relevance of the substantival-relational debate. Rynasiewicz contrasts the present state of the debate with its glorious past:

> What is remarkable about the substantival-relational debate is that, although it engaged natural philosophers from the seventeenth century into the nineteenth century and continues to be debated in academic philosophy, interest in the controversy on the part of twentieth century physicists has waned over the generations to virtually nil. (Rynasiewicz 1992, p. 588)

Meanwhile, Leeds questions the interest of interpretative work on GR, by contrasting the hole argument literature with an area of philosophy of physics which he regards as truly flourishing:

> There is an oddity here, it seems to me: for surely the philosophers of physics who work on these problems are the same men and women who, in another mood, are fond of comparing quantum mechanics with GTR, as the paradigm case of a theory which cries out for interpretation with the paradigm case of a theory which does not. (Leeds 1995, p. 428)

Discussion of the hole argument is often taken to be the epitome of irrelevant philosophy of physics. It is often held, implicitly or explicitly, that it is *obvious* that there is nothing to the argument, since no physicist would ever entertain for a minute the proposition that GR is an indeterministic theory. It is thus supposed to be somewhat of an embarrassment that

---

1970b, Earman 1970a, Earman and Friedman 1973, Sklar 1974, Sklar 1976, Glymour 1977). This discussion produced what I will be calling the received view. It is given its canonical presentation in Friedman 1983 and Earman 1989.

philosophers have wasted so much time on this argument—how do they expect physicists to take them seriously? Typically, partisans of this line of thought believe that the hole argument is predicated on some sort of simple mistake. Most spectacularly, it is claimed that it has nothing in particular to do with GR at all—rather it is an artifact of a certain misguided way of thinking about language, which Earman and Norton naively mistook for a bit of philosophy of physics:

> But now we are in a position to recognize that the hole argument has little to do with GTR *per se*. Rather, it is an instance of a problem that arises for *any* theory which posits nameable substances. (Maudlin 1989, p. 84)

> …the hole argument is really Putnam's argument restricted to spacetime theories. The key move is to realize that the gauge theorem is really about inscrutability of reference, not about indeterminacy of possible worlds. Thus, the gauge freedom in spacetime theories expresses a semantic fact rather than one about ontology. (Liu 1996 , pp. 1-2).

> Such permutation arguments have been exploited at length by W.V. Quine, Donald Davidson, and Hilary Putnam to argue that a hankering for absolute criteria of individuation leads to an inscrutability of reference. The hole argument is nothing more than an application of the same techniques to space-time theories. If it yields relationist or anti-realist conclusions, these are conclusions which apply globally to any ontology. The substantival-relational debate, however, was a local one over the status of space and time. (Rynasiewicz 1996a, p. 305)

It is further alleged that, having made this mistake, Earman and Norton have led other philosophers of space and time yet further astray by suggesting that the 'solution' to the hole argument is to be found in the furthest reaches of metaphysics:

> I think that issues about whether *this* spacetime point could in some other world have been over *there* are not really questions about the nature of spacetime points, or indeed about physics at all, they are questions about *situations* or *possible worlds*—philosophers' constructions so loosely connected with reality that we can consistently answer these questions in any way that we care to. And in fact it seems to me that this had begun to be the consensus about these questions until Earman seemed to breathe new life into them via the connection with determinism. (Leeds 1995, p. 436)

These philosophers of physics paint a bleak picture of the current state of the substantival-relational debate: having long ago lost its relevance to physics, it has recently degenerated into the worst sort of pedantic, confused, and eminently *philosophical* discussion.

Having been told that interest in the debate among physicists has waned to virtually nil, it is somewhat of a surprise to come across comments like the following in the physics literature:

> I would like to argue that the problem of quantum gravity is an aspect of a much older problem, that of how to construct a physical theory which could be a theory of an entire universe and not just a portion of

> one. This problem has a long history. It was, I believe, the basic issue behind the criticisms of Newtonian mechanics by Leibniz, Berkeley, and Mach. (Smolin 1991, p. 230)

It turns out that this comment is not atypical: many physicists who work on quantum gravity are interested in the substantival-relational debate, and believe that it is directly relevant to their research (see, e.g., Barbour 1995, Kuchar 1988, and Rovelli 1994). In particular, unlike Leeds, many physicists emphasize the importance of interpretative questions about GR—often motivated by the belief that differences of opinion about the technical and conceptual difficulties of quantum gravity can be traced to differences of opinion concerning the classical theory. Thus, Rovelli asserts that

> we believe that many discussions and disagreements on interpretational problems in the quantum domain (for instance the famous 'time issue') just reflect different but unexpressed interpretations of the *classical* theory. Thus, the subtleties raised by the attempts to quantize the theory force us to reconsider the problem of observability in the classical theory. (Rovelli 1991e, pp. 297-98)

Furthermore, far from dismissing the hole argument as resting on some sort of obvious mistake, many physicists see it as providing crucial insight into the physical content of GR. Thus, Isham uses a version of the hole argument to motivate an important claim about the observables of classical and quantum gravity:[4]

> …the diffeomorphism group moves points around. Invariance under such an active group of transformations robs the individual points of $\mathcal{M}$ of any fundamental ontological significance. … This is one aspect of the Einstein 'hole' argument that has featured in several recent expositions (Earman and Norton 1987, Stachel 1989). It is closely related to the question of what constitutes an *observable* in general relativity—a surprisingly contentious issue that has generated much debate over the years and which is of particular relevance to the problem of time in quantum gravity. In the present context, the natural objects that are manifestly Diff($\mathcal{M}$)-invariant are spacetime integrals… Thus the 'observables' of this type are intrinsically non-local. (§2.2 of Isham 1993)

Most surprisingly, one can even find physicists grappling with issues about the transworld identification of spacetime points:

> The basic principle of general relativity—as encompassed in the term "the principle of general covariance" (and also "principle of equivalence")—tell us that there is no natural way to identify the points of one spacetime with corresponding spacetime points of another. (Penrose 1996, p. 591)

It is clear that Penrose is led to this conclusion by the same concerns that motivate Rovelli and Isham (see, especially, p. 586). Interestingly, he goes on to attempt to formulate both

---

[4] Here, and throughout the dissertation, 'classical' is opposed to 'quantum.' 'Non-relativistic' is opposed to 'relativistic.'

criteria for approximate transworld identification, and measures of the accuracy of such approximations.

In short, a survey of the literature on quantum gravity reveals a very different picture of the relevance of philosophical work on the nature of space and time from that which is current among philosophers of physics. I do not mean to suggest that physicists are universally enthusiastic about the substantival-relational debate in general, or about the hole argument in particular. Nor, of course, are all philosophers of physics ill-disposed towards these topics. What *is* true is that there is a vocal subset of philosophers of physics who regard the substantival-relational debate as a waste of time, and who defend this view by invoking a consensus among physicists to this effect, although this consensus does not in fact exist. Clearly, then, philosophers of physics have a self-image problem. The major goal of this dissertation is broadly therapeutic: I want to make it possible for self-hating philosophers of physics to recognize the intellectual respectability of the substantival-relational debate.

This is my task in Part II, where I use considerations arising out of the hole argument to vindicate the physicists' claim that interpretative questions about ontological status of the spacetime points of classical GR are intimately related to the conceptual difficulties of quantum gravity. In Part I, I address a prior problem. In the second Part of the dissertation, I argue that certain conceptual and technical solutions to the problems of quantum gravity are associated with certain interpretations of the classical theory. In particular, some are associated with relationalism about the spacetime of GR, and some are associated with substantivalism. There are physicists who advocate both sorts of interpretation—Kuchar 1993a, for instance, is predicated on some sort of substantivalism, while Barbour 1994a, Rovelli 1994, and Smolin 1991 advocate varieties of relationalism. In the philosophical literature, however, there is a very broad consensus that relationalism about GR is untenable: the three most influential recent reviews of the substantival-relational debate—Earman 1989, Friedman 1983, and Maudlin 1993—are unanimous on this point;

and the deluge of responses to the hole argument seem to be, almost without exception, motivated by the belief that the lack of a viable alternative renders loyalty to substantivalism mandatory. In Part I of this dissertation my task is to examine the grounds which philosophers have offered for their rejection of relationalism. My motivation in undertaking this survey will be, looking ahead to Part II, to evaluate the interpretative possibilities for GR. It turns out, however, that the most convenient way to structure this discussion is to step backwards and first examine the arguments offered in discussion of the substantival-relational debate concerning non-relativistic physics. Thus only the last, and shortest, Chapter of Part I is directly concerned with the substantival-relational debate as applied to GR. It is preceded by two Chapters on the non-relativistic debate: Chapters 2 and 3 focus on the question of the existence of points of space and spacetime, respectively. My conclusion is that the received view is ill-founded, since no convincing reasons have been given to favor substantivalism over relationalism.

This allows us to proceed in Part II to look for connections between substantivalism and relationalism about GR and technical and conceptual maneuvers in quantum gravity. Now, given the ease with which they dismiss the relevance of the substantival-relational debate to physics, it is clear that philosophers of physics have not given sufficient attention to the problems of quantum gravity.[5] As a result, one finds a much greater discrepancy between the idioms that physicists and philosophers employ in their discussions of spacetime than one finds in the discussions of, e.g., quantum mechanics. Consequently, much of Part II is given over to recasting the hole argument in a form which is useful for discussion of the problems of the relation between the classical and quantum theories of gravity. Specifically, Chapter 5 is a brief expository introduction to the formalism of constrained Hamiltonian systems, which generalizes the familiar Hamiltonian framework of classical mechanics to allow for consideration of gauge degrees of freedom. In Chapter 6 I

[5] Hopefully this is already beginning to change, since papers by workers of quantum gravity have begun to turn up in the proceedings of philosophy conferences. See Page 1995, Sorkin 1995, and Unruh 1995.

discuss the general connection between determinism and ontology, using examples of constrained Hamiltonian systems to show that questions of determinism are dependent on questions of interpretation. The hole argument makes its appearance in Chapter 7. There I employ the gauge-theoretic structure of GR in classifying the strategies for interpreting the theory. I show that under many interpretations—including the most plausible form of substantivalism—GR is an indeterministic theory. Some popular substantivalist responses to the hole argument are considered—and found wanting—in the Appendix.

In Chapter 8, I at last attempt to make good my promise to find a connection between the problems of quantum gravity and interpretations of GR. I begin with examples from quantum mechanics and quantum field theory which illustrate the way in which quantization can raise questions of interpretation. In the case of quantum gravity, I argue that certain strategies for interpreting GR are naturally associated with certain strategies for constructing a theory of quantum gravity. Thus reasons for accepting a given interpretative stance towards the classical theory will count as a reason to endorse a certain approach to quantum gravity—and vice versa. In particular, I argue that the sort of substantivalist interpretation under which GR is indeterministic suggests (and is suggested by) a certain approach to defining the observables of quantum gravity. Similarly, there is a suggestion about how we can understand the role of time in quantum gravity which is closely related to a variety of relationalism according to which there is after all a preferred notion of simultaneity in GR. Thus our beliefs about the ontology of the classical theory influence, and are influenced by, the details of the emerging quantum theory of gravity.

Finally, in the Afterword I offer a summary of the argument and outline some questions for further research. My chief conclusion is that the interpretative interplay which exists between classical and quantum gravity means that the substantival-relational debate is of more importance and interest than it has been given credit for in recent philosophical discussion.

The remainder of this first Chapter is given over to an attempt to place the substantival-relational debate in a philosophical context. §2 distinguishes the 'relationalism' of the substantival-relational debate from a couple of its close cousins. In §3 I argue that the debate has little if anything to do with the debate over realism concerning scientific entities in general. §4 contains an attempt to locate the debate within the broader program of interpreting physical theories, and some discussion of the interpretative methodology which I employ in the body of the dissertation.

2. The substantival-relational debate is part of a knot of difficult questions about the nature of space, time, and motion which are often grouped together under the general heading of questions about 'relationalism.' Earman 1970b and Horwich 1978 provide quite thorough catalogs of the various senses of relationalism and the connections between them. Here it will suffice to distinguish the substantival-relational debate from two closely related topics.

The first is the intrinsic-conventional debate. The question here is whether the geometric structure of spacetime is intrinsic and is disclosed by our techniques of measurement, or is indeterminate and is fixed by a conventional choice of a standard of geometric measurement. Clearly this question of the nature of the relations between the parts of spacetime is closely related to the question of the nature of the existence of the parts themselves. So one might worry that settling the intrinsic-conventional debate would settle the substantival-relational debate. But notice that the sort of considerations that bear on the intrinsic-conventional debate are in general quite irrelevant to the substantival-relational debate. For instance: conventionalism is incompatible with the reduction of the properties of matter to properties of spacetime à la geometrodynamics, since if spacetime is all that there is then its geometric structure must be intrinsic (see Grünbaum 1973, §22.3 and Grünbaum 1977, p. 318); but the viability of geometrodynamics is compatible with both substantivalism and relationalism. In fact, I know of no arguments which demonstrate that taking a position on one of these questions constrains one's choice of position on the other question. Furthermore, it is clear that a great deal can be said either without delving into the intricacies of the other (compare, e.g., Grünbaum 1973 with Earman 1989). So I feel quite safe in assuming that the two questions can be treated as being logically independent.

The absolute-relative debate concerning the nature of motion, on the other hand, cannot be neglected in a discussion of the substantival-relational debate. Since the Seventeenth century it has seemed plausible that there is a close connection between the two issues: space is relational iff motion is relative (see §2.2). Today, of course, the situation looks considerably more complex than it once did, since we now have to worry about spacetime rather than about space, and to consider a variety of theories of motion. Nonetheless, although an evaluation of the senses in which motion is or is not relative in various physical theories is well beyond the scope of this dissertation, it is impossible to evaluate the virtues of substantivalism and relationalism without touching on the connection between absolute motion and substantivalism. This is particularly true in the special case, to be considered in Chapter 2, of the substantival-relational debate as applied to space in the context of non-relativistic physics.

3. The substantival-relational debate may be characterized as an argument over the ontological status of certain entities—the points of space or spacetime—which are postulated by scientific theories. So it might at first seem plausible that the moves in this debate are going to be intimately related to familiar moves in the general debate concerning scientific realism. I do not believe that this is correct, since I see scientific realism as being a doctrine about what it means for the entities postulated by science to exist, whereas I see the substantival-relational debate as being a debate about whether or not the existence of spacetime points should in fact be postulated.

It is helpful to have a clear characterization of scientific realism on the table before embarking on a discussion of this notoriously slippery topic. I like Boyd's influential characterization, which identifies scientific realism with the conjunction of the following four theses:

1. Theoretical terms in scientific theories … should be thought of as putatively referring expressions; that is, scientific theories should be interpreted "realistically."
2. Scientific theories, interpreted realistically, are confirmable and in fact often are confirmed as approximately true by ordinary scientific evidence interpreted in accordance with ordinary methodological standards.
3. The historical progress of the mature sciences is largely a matter of successively more accurate approximations of the truth about both observable and unobservable phenomena. Later theories typically build upon the (observational and theoretical) knowledge embodied in previous theories.
4. The reality which scientific theories describe is largely independent of our thoughts or theoretical commitments. (Boyd 1984, pp. 41-42)

For Boyd, anti-realists fall into two sorts: empiricists who reject the first two theses and are chary about the third; and constructivists who reject the fourth, and place an appropriate gloss on the first three.

Now, there are at least three ways in which participants in the substantival-relational debate understand its relation to the debate on scientific realism:

(1) Some philosophers argue that substantivalism is just a sub-thesis of scientific realism. Thus we have Butterfield professing that:

> scientific realism holds that one is committed to believing in the existence of those entities that are ineliminably referred to or quantified over by one's best scientific theories. And our best spacetime theories are almost always presented as quantifying over spacetime points…. As an aspiring scientific realist, I find … substantivalism attractive. (Butterfield 1989b, p. 2)

This is just the familiar Quinean line that one should read one's ontology off from the preferred formulation of one's scientific theory. If we accept this as the sine qua non of realism, then it follows that one's position on the substantival-relational question is determined by one's position on the question of scientific realism (see also Earman and Friedman 1973, pp. 329 and 358; Hoefer 1996, p. 23; and Nerlich 1993).

(2) van Fraassen 1995 seems to accept Butterfield's contention that realists should adopt a Quinean approach to scientific ontology, but he recognizes that this leaves them considerable room for maneuver. Thus he reads Earman 1989 as presenting the substantival-relational debate as a squabble between realists concerning whether or not the preferred formulation of GR *does* quantify over spacetime points (p. 144). van Fraassen believes that relationalists, on the other hand, should reject the Quinean criterion (see

especially p. 145). He approvingly characterizes Friedman 1983 as presenting the substantival-relational debate as a confrontation between a Butterfield-style substantivalist and a constructive empiricist relationalist. Presumably this is because Friedman, like van Fraassen himself, sets up the dispute between realists and anti-realists as a question of whether by restricting their ontology to a subset of the ontology of the former, the latter increase or diminish the level of confirmation of their theory (Friedman 1983 and van Fraassen 1980). Thus for van Fraassen, the realist may be either a substantivalist or a relationalist, but the (constructive empiricist) anti-realist is committed to believing only in the observable, and hence must be a relationalist.

(3) Earman express the hope that, although there is no straightforward connection between the two debates, there is still hope that progress in one will lead to progress in the other:

> the absolute-relational controversy taps some of the most fundamental concerns in the foundations of physics, metaphysics, and scientific epistemology. This is a cause for both despair and hope: despair because if the absolute-relational controversy cannot be resolved without first settling the big questions of metaphysics and epistemology, it is not likely to be resolved, and hope because a way of making progress on the absolute-relational controversy can lead to progress on the big questions.[6]

In order to describe the nature of my disagreement with these three theses, it may be helpful to borrow Hacking's "gruesome analogy" (Hacking 1983, pp. 95-96). The idea is to compare disputes over realism with various sorts of military conflict. Thus, wholesale disputes about the nature of physical reality, such as the debate between realists and idealists over the mind-independence of reality, are compared to a total war in which the great powers directly engage one another. In such an all-out conflict, the stakes are very high—the winner determines what it means to be real. According to Hacking, the debate over scientific realism is, on the other hand, a 'mere' colonial war: the nature of the real is not at stake; scientific realists merely want to annex new territory to the empire of the real, while anti-realists want to prevent this. Notice that each of the three theses above is explicitly concerned with only the first three of Boyd's theses—i.e., none of them engages more global concerns about the

[6] Earman 1989, p. 2. Here Earman's 'absolute-relational controversy' includes both my 'substantival-relational debate' and my 'absolute-relative debate.'

mind-independence of physical reality. Thus, they cast the substantival-relational debate not as part of the total war, but as a small—but bloody—incident in a colonial war.

I don't think that the substantival-relational debate is part of the shooting war at all. It is more like a squabble at the headquarters of one or the other army: "Is this blotch on the map a hill or a stain?" Consider the case of the vector potential in classical electromagnetism (this example will recur throughout the dissertation, for reasons which will become obvious in Part II). Prior to any questions about the proper attitude to take towards the entities postulated by science—let alone questions about the mind-independence of reality—one has to understand the status of the vector potential in the classical theory. The position which was usually adopted in the Nineteenth century was that it was a mere mathematical fiction, while the electric and magnetic field vectors corresponded to elements of physical reality. But one can imagine a debate over the status of the vector potential, with one faction insisting that it was just as real as the electric field, the other insisting that it was a mere representational artifact. Once this debate is settled we have an interpretation of electromagnetism. We can *then* go on to ask what attitude to take towards the entities postulated by electromagnetism. Just so with spacetime points—the question as to whether or not they are part of the ontology or GR is prior to, and independent of, the scientific realism debate. In the martial analogy: the substantival-relational debate is not part of the battle, it is part of an internecine dispute about whether or not spacetime is even an *objective* .

Thus I find both Butterfield's and van Fraassen's positions to be unnecessarily constraining: I do not see why realists (anti-realists) should have any particular stake in the substantival-relational debate, since it shouldn't matter to them whether this or that putative theoretical entity makes it on to their short list of genuine theoretical entities, so long as they are free to take such entities seriously (with a grain of salt). Similarly, I find Earman's position unrealistically optimistic. It is difficult to see how considerations bearing on the question of whether one or another specific entity is a respectable scientific entity can be parlayed into progress on the question of the ontological status of scientific entities in general. This is because on most accounts of realism and anti-realism, these two questions are logically independent.

The exception is the deflationary attitude towards realism advocated in Cartwright 1983, Fine 1984b, Fine 1984a, and Hacking 1983. Their idea is that the only interesting question in the neighborhood of scientific realism is: what is it that changes a putative entity from something whose existence is a possibility that scientists entertain into something whose existence they presume? And now we can expect to make some progress on the larger issues by looking at special cases: electrons, gravity waves, weak neutral currents…

I very much doubt that this is the kind of progress that Earman is hoping for. In any case, it is questionable whether even dedicated deflators will be optimistic about the prospects. After all, one of the most characteristic features of the deflationary program is its emphasis on the 'local': there is no essence of science, so one should not look for invariants across time, space, or discipline (see especially Fine 1984a and Fine 1996). For the deflator, there is a constant danger of being misled by a mode of inference which is applied outside of its proper context. I would claim that anyone (and not just deflators) should be skeptical about the prospects of drawing any meta-lessons from the substantival-relational debate. If spacetime points are entities, then they are peculiar entities. They don't, for instance, possess causal efficacy, which many consider to be the hallmark of a respectable physical entity (see Cartwright 1983 and Salmon 1984; also see Nerlich 1979 for skepticism about whether spacetime as a whole deserves to be considered to be causally efficacious). And so one should suspect that whatever arguments bear on the existence of spacetime points are of a quite special nature. There is little hope that "a way of making progress on the absolute-relational controversy can lead to progress on the big questions."

4. The invention, articulation, praise, and criticism of interpretations of physical theories are some of the central tasks of philosophy of physics. This is, of course, most clear in work in the foundations of quantum mechanics. Redhead offers a quite typical characterization of an interpretation of quantum mechanics: "some account of the nature of the external world and/or our epistemological relation to it that serves to *explain* how it is that the statistical regularities predicted by the formalism … come out the way that they do" (Redhead 1987, p. 44). And I believe that he is also typical in going on to speak of all of physics in claiming that "theories which lack an interpretation … simply do not contribute to our *understanding* of the natural world." But, of course, no known theory is both fully relativistic and fully quantum. It follows that all of the theories which philosophers of physics are so interested in interpreting are, without exception, *false*. This is a fact which is seldom mentioned, and whose significance is very rarely discussed. In this section I would like to sketch a framework for understanding the value of interpreting false theories.

Although there *is* something strange about interpreting a false theory, such work can always be defended as contributing to our understanding of the content of the theory. But is it true, as Redhead claims, that this can contribute to our understanding of the *world*? I think that it is helpful here to shift perspective slightly, and to begin by considering the set of our false theories as a structured network, rather than as a set of isolated theories.

At the present time we have a number of theories on the books: classical mechanics, statistical mechanics, electromagnetism, quantum mechanics, quantum statistical mechanics, quantum field theory, special relativity, general relativity, quantum field theory on a curved background, and (partially) quantum gravity, and so on.[7] These theories tell us about very different worlds. Some are populated by particles, some by fields. In some the spatiotemporal structure is an unchanging backdrop, in some it is an active and changing participant. It is an important fact that despite the nontrivial overlap between the domains of applicability of pairs of these apparently incompatible theories we find that peaceful co-existence rather than competition is the rule. Our world is somehow represented by each of these diverse theories.

[7] See p. 6 of Guillemin and Sternberg 1984 for a diagram of the relations between six theories of light.

So we have a network of independent theories. This network is often described as a hierarchy, the idea being that some theories are more fundamental than others. I think that a web or a lattice would be a more appropriate metaphor here, since theories often have more than one limit. For instance, special relativity is the curvature→0 limit of GR, while the Newton-Cartan theory is its c→∞ limit (Malament 1986). It is still possible to speak of one theory being more fundamental than another, so long as we don't make the mistake of assuming that 'more fundamental' gives us a linear ordering of the class of theories. (Which is more fundamental, quantum statistical mechanics or electromagnetism?)

Each of our theories is thought to be empirically adequate within its own domain of applicability, but to share a part of this domain with another theory which is also supposed to save the phenomena. Given this situation, it seems essential to demand for every pair of overlapping theories an assurance that they really do agree on their shared domains. Thus I read the correspondence principle as requiring that quantum mechanics be able to account for the empirical adequacy of classical mechanics. It is important to emphasize that one cannot expect too much of the correspondence principle and its generalizations. As Rohrlich argues, it is not in general possible to show that as one takes a classical or non-relativistic limit the ontology of a given theory goes over into the ontology of the limiting theory (see Batterman 1995, Berry 1994, and Rohrlich 1988). And this should not bother us if we keep in mind that both the given theory and the limit theory are false. Similarly, it is important to keep in mind that the correspondence principle requires only that the claim that each of the theories is empirically adequate be consistent, and not that the theories be empirically equivalent. This makes it possible, for instance, maintain the truth of the correspondence principle, even in the face of the existence of periodic quantizations of chaotic classical systems (Belot and Earman 1996a).

Thus each of our physical theories is part of a network theories which stand in subtle relations to one another. Each theory contributes to our understanding of the world not only in virtue of its own content, but also because of the relations in which it stands to other theories.

Now let's turn to the question the utility of interpretation. Since any given theory is false, knowing its interpretation does not tell us what the world is like. But it is tempting to think that the fact that a given theory is empirically adequate in its proper domain gives us *some* information about the world—and thus that an interpretation of the theory, which tells what the world could be like were the theory true, tells us something about what the world is actually like. And it is natural to assume that more fundamental theories tell us *more* about the world. I believe that there is some truth to this line of thought. But carrying it to its logical conclusion suggests that much of the interpretative work that philosophers of physics pursue is misguided:

> it is worth noting that some recent developments in high-energy physics indicate that space-time may not be a fundamental property of nature, but only an approximation arising in the limit of high complexity. If this should turn out to be so, then many of the foundational studies of space-time by philosophers would seem misplaced. That is, one should be careful where he looks to choose foundational problems. For a start, he had better make certain that the theory or concept is a fundamental one, and not just some approximation. (Cushing 1990, p. 33)

Cushing's view seems to be that a foundational/interpretative problem is interesting only if it, or a cognate, arises in a true theory, because only then is it a problem about the world rather than a problem about a false theory. He seems to take these hints from high-energy physics as a sign that spacetime will cease to be part of the basic language of physics as we formulate more and more fundamental theories, and concludes that it is a mistake to do interpretive work on spacetime theories.[8]

Given the discussion of §1 above, it should be obvious what my answer to Cushing will be: I believe that interpretative questions can contribute to our understanding of the relations between theories. This should not come as a surprise since, for instance, it is well-known that the processes of quantization and taking a classical limit are not purely formal,

---

[8] One might add that recent work on quantum gravity suggests that GR itself contains the seeds of the destruction of spacetime, in the sense that the application of the rules of canonical quantization to GR apparently leads to a theory in which spacetime is not fundamental; see Ashtekar 1995. For other approaches to discrete spacetime, see Madore 1996 and Sorkin 1995, and the references therein.

but depend to some extent on interpretation. In Chapter 8, I argue in detail that in the case of the relation between GR and quantum gravity we find that there is an interrelation between interpretations of the classical theory and various proposed solutions to the outstanding conceptual problems of quantum gravity. I also present examples of similar intertheoretic interpretative dynamics which result from quantizing ordinary finite dimensional systems and classical field theories on fixed spacetime backgrounds.

I believe that this fact provides a very respectable reason for taking the project of interpreting GR seriously, despite the fact that the spacetime which is the subject matter of that theory may not be present in more fundamental theories. We see that an interpretation of a given theory, in addition to contributing to our understanding of the world directly by describing a possible world more or less close to our own, may also help us to understand—or even to formulate—other nearby theories. That is, in the absence of a true theory, all of our information about the world is contained in our network of overlapping false theories, and there are two ways in which interpretation helps us to understand what this network tells us about the world: (i) by revealing the content each member of the network; (ii) by articulating the structure of the network through clarification of the relations that hold between its members. (Note that this is a picture which realists and the anti-realists should be able to agree on—since, as argued in the previous section, all hands should be interested in isolating the content of science, before going on to argue about what attitude we should take towards this content.)

Before closing this discussion of interpretation, I would like to mention three features of the interpretative methodology that I will employ in the body of the dissertation. The first is that I take it for granted that theories do not interpret themselves. There is a widespread tendency to assume that it is easy to read off the interpretation of a theory from its most perspicuous formulation. Most philosophers of physics do, of course, realize that this cannot always be done—no-one would claim that it is *obvious* how one should interpret quantum mechanics. But it is not uncommon to claim that certain theories wear their correct interpretations on their sleeves (see especially Leeds 1995, p. 436). I do, of course, admit that some theories have literal interpretations, in which there is a bijection between the mathematical structure of the models of the theory and the ontology of the possible worlds represented by those models. But I insist on the fact that every theory admits of many interpretations. Some of these may not be obvious, but they may be interesting nonetheless. For instance, as we will see in Chapter 6, any gauge theory is indeterministic under its literal interpretation, so that one tends to prefer non-literal interpretations when dealing with such theories.

Second, up until Chapter 8 I will be exclusively concerned with the extent to which intratheoretic considerations serve to settle interpretative questions. At each stage I will be considering a single theory—Newtonian physics in Chapters 2 and 3, GR in Chapters 4 and 7—and will inquire after resources internal to the problematic of the theory itself which mitigate in favor of one interpretation or another. I will not enter into questions such as whether the merits of substantivalism in the non-relativistic regime makes it more attractive as a doctrine about the spacetime of GR (see Friedman 1983 and Maudlin 1993 for this sort of approach). In Chapter 8, I will lift this prohibition and will look at the interpretative consequences of intertheoretic relations.

Finally, I reserve the right to cut off discussion when I feel that the issue at hand reduces to a clash of metaphysical intuitions. I will not, for instance, allow it to count against relationalism that it is not, while substantivalism is, compatible with a thorough-going nominalism about mathematical entities (see Field 1980 and Field 1985). At its best, the substantival-relational debate has always derived part of its excitement from its relation to issues of physics. Thus, I think it is quite appropriate to focus on the criteria for comparing substantivalism and relationalism which arise out of physics, and to be satisfied with an evaluation of their strengths which holds only upto issues of metaphysics.

PART I
Rehabilitating Relationalism

CHAPTER 2
The Non-Relativistic Debate: Space

1. My task in this first Part of the dissertation is to survey the literature on the substantival-relational debate, and to argue against the received view that classical physics provides conclusive grounds for ruling out relationalism as an interpretative strategy. This sets up the discussion of Part II as a genuine horse race between substantivalism and relationalism.

Ultimately, I will be interested in assessing the interpretative possibilities for GR. But I find it convenient to begin by surveying the standard arguments concerning the supposed superiority of substantivalism over relationalism in the non-relativistic context. This allows me to deal with the arguments individually in a relatively tame setting before turning to GR where they must be faced as a corporate body. I carry out this task in this Chapter and the next, concluding that each of the standard arguments in ineffective. In Chapter 4 I carry over these conclusions to the case of GR. Throughout this first Part, I restrict my attention to intratheoretic concerns: in this Chapter and the next, I ask what considerations arising out of Newtonian physics could lead us to favor one interpretative approach over the other; in Chapter 4 I do the same for GR. For the purposes at hand, 'Newtonian physics' will for the most part mean the Newtonian theory of gravitating point particles—although fields will make a brief appearance in this Chapter. The idea, however, is to let this fragment of classical mechanics stand in for all of pre-relativistic physics. In what follows, I will demand that the interpretative strategies under consideration be capable of handing the full range of non-relativistic physics. This demand is to some extent a convenience—I could, instead, require merely that the interpretative strategies be appropriate for some empirically adequate fragment of Newtonian physics. We will see in the final sections of this Chapter that adopting this weaker condition would have some interesting consequences. Nonetheless, I choose to utilize the stronger condition, which has been traditionally employed in philosophical discussion of the substantival-relational debate. In any case, this choice makes no material difference to the discussion of Part II.

Originally, of course, substantivalism and relationalism were doctrines about the nature of the existence of space. In recent decades, however, it has become increasingly standard to view substantivalism and relationalism as being doctrines about the nature of the existence of spacetime. Thus the traditional disagreement concerning the existence of the points of space has been replaced by a disagreement concerning the points of spacetime. So we now have two versions of each of substantivalism and relationalism.[9] In Chapter 3, I examine arguments which are supposed to demonstrate the superiority of substantivalism about spacetime over relationalism about spacetime. My task in this Chapter is to investigate the viability of three dimensional interpretations of non-relativistic physics.

The structure is as follows. In the next, section I provide a brief summary of the doctrines of Newton, Clarke, Leibniz, and Locke. The goal is to motivate the notions of substantivalism and relationalism about space which are in play in the contemporary debate, and to present some of the details of the historical debate which remain relevant. I then explicate the notions of substantivalism and relationalism about space which I will be employing, and defend relationalism against the accusation of incoherence. Finally, I turn to the two most important reasons which are given for rejecting relationalism about space: that it is inappropriate for field theories, and that it is incapable of handling inertial effects.

2.

(i) Newton.

[9] Where it is important to distinguish between the different notion I will use terms such as 'spacetime substantivalism' and 'three dimensional interpretations.'

Newton's clearest pronouncements concerning the nature of space are contained in his manuscript *De Gravitatione*.[10] Much of this essay is an attack on Descartes' physics, which is supposed to be founded on a relational theory of space and a relative theory of motion. This theory is presented in *Principles of Philosophy*, where it is nicely summarized in §II.13 (Descartes 1967):

> The terms 'place' and 'space', then do not signify anything different from the body which is said to be in a place; they merely refer to its size, shape and position relative to other bodies. To determine the position, we have to look at various other bodies which we regard as immobile; and in relation to different bodies we say that the same thing is both changing and not changing its place at the same time.

In *De Gravitatione,* Newton is concerned to argue that Descartes' account is incoherent:

> that the absurdity of this position may be disclosed in full measure, I say that thence it follows that a moving body has no determinate velocity and no definite line in which it moves. And, what is worse, that the velocity of a body moving without resistance cannot be said to be uniform, nor the line said to be straight in which its motion is accomplished. On the contrary, there cannot be motion since there can be no motion without a certain velocity and determination. (p. 129)

The heart of Newton's argument is his claim that "Truly there are no bodies in the world whose relative positions remain unchanged with the passage of time…" (p. 130). From this it should follow that for a Cartesian "there is no basis from which we can at present pick out a place which was in the past, or say that such a place is any longer discoverable in nature." The conclusion:

> it is clear that if one follows the Cartesian doctrine, not even God himself could define the past position of any moving body accurately and geometrically now that a fresh state of affairs prevails, since in fact, due to the changed positions of the bodies, the place does not exist in nature any longer.

From this, it follows that Descartes is not entitled to the concept of velocity:

> Now as it is impossible to pick out the place in which a motion began (that is, the beginning of the space passed over), for this place no longer exists after the motion is completed, so the space passed over, having no beginning, can have no length; and hence, since velocity depends upon the distance passed over in a given time, it follows that a moving body can have no velocity, just as I wished to prove at first.

A fortiori, Descartes is not entitled to the law of inertia, and his entire mechanics must collapse.

Newton believes that in demolishing the Cartesian account he has also shown that "it is necessary that the definition of places, and hence of local motion, be referred to some

---

[10] First published in Hall and Hall 1962. The Halls speculate that this is an early essay, dating from 1669. Stein offers some reasons to believe that it may be a much later work (Stein 1970, fn. 11).

motionless thing such as extension alone or space in so far as it is seen to be distinct from bodies" (p. 131). This justifies his own approach to the nature of space and motion. He begins by telling us that the term 'space' is "too well known to be susceptible of definition by other words" (p. 122). But his definitions of place and motion reveal quite a bit about his conception of space: "place is a part of space that something fills evenly"; "Motion is change of place." As he says: "I suppose in these definitions that space is distinct from body, and … I determine that motion is with respect to the parts of space, and not with respect to the position of neighboring bodies" (p. 123).

Newton later clarifies the sense in which space exists. Following Gassendi, he maintains that "it has its own manner of existence which fits neither substances nor accidents" (p. 132; for the similarity to Gassendi, see §§8.4.b and 8.4.l of Grant 1981).[11] On the one hand it is not a substance. It fails the classic test of substantiality: it depends on something else for its existence (namely, God). Furthermore, it is not active in the way that paradigm cases of substance (body and mind) are. On the other hand it is not an accident: we believe that it exists in the absence of body ("as when we imagine spaces outside the world, or places empty of body"), so "we cannot believe that it would perish with the body if God should annihilate a body, it follows that [extension] does not exist as an accident inherent in some substance."[12] Some of Newton's contemporaries, such as Pierre Bayle, argue that extension does not exist at all (see the article "Zeno of Elea" in his *Philosophical Dictionary*, Bayle 1965). Newton is careful to distinguish himself from such paradox-mongers: "And much less let it be said to be nothing, since it is rather something, than an accident, and approaches more nearly to the nature of substance."

Newton goes on to give a list of the properties of space. For our purposes, the most important are: that it consists of parts (pp. 132-33); that "The positions, distances, and local

[11] This long paragraph is nominally concerned with extension rather than space. But I think it is clear that Newton is using the terms interchangeably. See especially his list of the properties of our idea of extension: sometimes we are given what is supposed to be a property of space and no mention at all is made of extension (pp. 132-8).
[12] Note that, at best, this shows only that extension is not an accident of body. This leaves open the possibility that it is an accident of God, as Clarke seems to have believed.

motions of bodies are referred to the parts of space" (p. 137); and that "The parts of space are motionless" (p. 136). Thus, aside from the question of whether or not it is a substance, it seems that Newton's view of space is relatively straightforward: it is an existing thing whose parts maintain their identity through time.[13] Furthermore, we are required to posit its existence if we want to employ the notion of velocity in our physics.

(ii) Leibniz-Clarke.

The Leibniz-Clarke Correspondence of 1715-16 is the most comprehensive and accessible of Leibniz's published pronouncements on the nature of space and time. It also poses a serious interpretative problem for the student of the philosophy of space and time. Leibniz describes the correspondence to Bernoulli by saying that "You perhaps know that I am at present engaged in a philosophical quarrel with Newton or, what amounts to the same thing, with his defender Clarke" (Alexander 1956, p. 189). Unfortunately, it no longer seems so obvious that it *does* amount to the same thing, since historians disagree about the

[13] It is not clear whether Newton would permit the substitution of 'points' here for 'parts.' On the one hand, he argues in an early notebook that space is made up of extended indivisibles which are explicitly contrasted with points ("[a] least extension is infinitely larger y^{n} a point & therefore can conteine it"; see §2 of McGuire 1982, especially pp. 150 and 154). Apparently, he was motivated here by the desire to give an account of the metric nature of space in terms of the nature of its parts; and by the common belief that nothing composed of points could be extended (on this point see, e.g., Hume in the *Treatise*: "the system of mathematical points is absurd; and that system is absurd, because a mathematical point is a non-entity, and consequently can never by its conjunction with others form a real existence"; Hume 1978, p. 40). Thus Grünbaum is perhaps too hasty in contrasting Newton's views on the intrinsic-conventional debate with Riemann's, since it seems that the early Newton was well aware of the metric amorphousness of the continuum (Grünbaum 1973, pp. 8-9). In this connection, it is interesting to note that Sorkin 1995 mentions that one of his reasons for becoming interested in approaches to quantum gravity in which spacetime is fundamentally discrete was the feeling that such structures support a more satisfying account of the metric nature of spacetime.

On the other hand, a manuscript from the early 1690's (first published in McGuire 1978) suggests that Newton was considering amending the Scholium on space and time of the *Principia* to explicitly mention that space is composed of points: "Space is not compounded of aggregated parts since there is no least in it, no small or great or greatest, nor are there more parts in the totality of space than there are in any place which the very least body of all occupies. In each of its points it is like itself and uniform nor does it truly have parts other than mathematical points, that is everywhere infinite in number and nothing in magnitude" (p. 117). These emendations were never made, however, so it is not clear which view we should take to represent Newton's considered opinion.

extent to which Newton was involved in the dispute. Koyré and Cohen 1962, on the one hand, argue that he drafted parts of Clarke's letters. Hall, on the other, is very skeptical:

> this is not to say that Newton actually guided Clarke's pen or stood behind his shoulder as he wrote; of this we have no evidence at all, and the declaration of the two most recent students of the matter that "there is no doubt that Newton took part in the fight between Leibniz an Clarke," literally read, goes beyond what was doubtless their intention to express. Certainly Newton was in Clarke's corner, but he did not on more than an isolated occasion at most put lead in Clarke's gloves. (Hall 1980, p. 220)

I will adopt the safest course, and assume that Clarke is always speaking for himself.

Whether or not it provides an accurate portrait of Newton's views, the Correspondence is still the clearest classical discussion of the substantival-relational debate. As such, it provides an invaluable anchor for contemporary discussions. This is not, however, because either Leibniz or Clarke offers a particularly clear account of the nature of space.

From Clarke we hear that "Space is not a being, an eternal and infinite being, but a property or a consequence of a being infinite and eternal" (Alexander 1956, p. 31). So it seems that Clarke, unlike Newton, regards space as an attribute. But, as noted above, Newton's (then unpublished) reasons for denying that space is an attribute do not seem to rule out Clarke's position. So there is room to wonder how far apart Newton and Clarke really are, especially given Newton's claim that space is "as it were an emmanent effect of God" (Hall and Hall 1962, p. 132; see also Queries 28 and 31 of the *Opticks* (Newton 1952), where he speaks of space as the sensorium of God). This question is further obscured by the fact that Des Maiseaux's preface to the Correspondence includes a qualification of Clarke's views, supposedly written by Clarke himself, but attributed to Newton by Koyré and Cohen:

> he does not claim to take the term *quality* or *property* in the sense that they are taken by those who discuss logic or metaphysics when they apply them to matter; but that by this name he means only that space and duration are modes of existence of the Substance which is really necessary, and substantially omnipresent and eternal. This existence is neither a substance nor a quality nor a property; but is the existence of a Substance with all its attributes, all its qualities, all its properties; and place and duration are modes of this existence of such a kind that one cannot reject them without rejecting existence itself. When we speak of things which do not fall under the senses it is difficult to speak without using figurative expressions. (p. xxix)

In his *New Essays on Human Understanding,* Leibniz appears to join Newton in affirming the existence of space, while denying that it is either a substance or an attribute.[14]

PHIL. **§17**. If anyone asks "whether… space void of body, be substance or accident, I shall readily answer, I know" nothing about it.
THEO. I have reason to fear being accused of vanity in trying to settle what you, sir, admit you do not know. But there are grounds for thinking that you know more about it than you say or believe that you do. Some people have thought that God is the place of objects: Lessius and M. Guericke, if I am not mistaken, held this view; but it makes place involve something over and above what we attribute to space, to which we deny any agency. Thus viewed space is no more substance than time is, and if it has parts it cannot be God. It is a relationship: an order, not only among existents, but also among possibles as though they existed. But its truth and reality are grounded in God, like all eternal truths.
PHIL. I am not far from your view. You know that passage in St. Paul which says that in God we live, move and have our being. So that depending on how one looks at the matter, one could say that space is God or that it is only an order or relation.
THEO. Then the best way of putting it is that space is an order but that God is its source. (Leibniz 1981, pp. 149-50)

This relational account of the nature of space is amplified in the Correspondence:

As for my own opinion, I have said more than once, that I hold space to be something merely relative, as time is; that I hold it to be an order of coexistences, as time is an order of successions. For space denotes, in terms of possibility, an order of things which exist at the same time, considered as existing together; without enquiring into the their manner of existing. And when many things are seen together, one perceives that order of things among themselves. (pp. 25-26)

Thus far, it is difficult to make sense of the differences of opinion between Newton, Clarke, and Leibniz. All seem to agree—contra Bayle—that space exists, and that it is not easy to explicate this existence within the framework of traditional metaphysical categories. It is natural to worry that there is no real difference of opinion here, merely different descriptions of the same sort of existence (see Stein 1967, pp. 193-94; and §2 of DiSalle 1994). The lasting contribution of the Leibniz-Clarke Correspondence is to dispel this worry. Leibniz and Clarke manage to agree on a criterion that distinguishes the relationalism of Leibniz, on the one hand, from the so-called substantivalism of Clarke and Newton, on the other. Furthermore, this criterion is *physical*: space is relational iff there is no absolute position or motion. In what follows, it will be helpful, as well as more accurate, to speak of Newton and Clarke as absolutists about space, rather than as substantivalists. Roughly, this

---

[14] About a year before his death, Leibniz wrote a letter to Conti in which he commented that "Space is something; but, like time, it is a general order of things. Space is the order of coexistents, and time is the order of successive existents. They are true things, but are ideal, like numbers" (p. 446 of Vol. III of Dutens 1768); my translation. Note that the translation appearing on p. 185 of Alexander 1956 omits the crucial first sentence. See §1.4 of Earman 1989 for a discussion of the role of 'ideal' in Leibniz's thought about the nature of space.

will mean that Newton and Clarke (contra Leibniz) take the parts of space to be beings which maintain their identity over time, independently of the distribution of matter.

Let's say that position is absolute if it depends, as in Newton's account, on the parts of space rather than on material bodies. This notion enters the Correspondence in the course of an attempt by Clarke to discredit Leibniz's reading of the Principle of Sufficient Reason (p. 20). His idea is to show that there are cases where the will of God alone must be a sufficient reason for something: "For instance: why this particular matter should be created in one particular place and not in another particular place." Clarke seems to have stumbled unwittingly onto one of the major differences of opinion between himself and Leibniz concerning the nature of space.

Leibniz responds with one of the greatest bluffs of all time: "I have many demonstrations to confute the fancy of those who take space to be a substance, or at least an absolute being. But I shall use, at the present, one demonstration, which the author here gives me occasion to insist upon" (p. 26). Essentially, Leibniz reiterates Clarke's point, but claims that it constitutes a reductio of the absoluteness of space. If space were an "absolute being" then Clarke would be correct: God would have no reason to create the material universe in one region of space rather than another.

> [W]ithout the thing placed in it, one point of space does not absolutely differ in any respect whatsoever from another point of space. Now from hence it follows, (supposing space to be something in itself, besides the order of bodies among themselves) that 'tis impossible there should be a reason, why God, preserving the same situations of bodies among themselves, should have placed them in space after one peculiar manner and not another…

This presupposes that the absoluteness of space implies the absoluteness of position. Now by the Principle of Sufficient Reason, we have a contradiction, and it follows that space is not absolute, in which case Clarke's shift argument is ineffectual. On the other hand,

> if space is nothing else, but that order or relation; and is nothing at all without bodies but the possibility of placing them; then those two states, the one such as it now is, the other supposed to be quite the contrary way, would not at all differ from one another. Their difference therefore is only to be found in the chimerical supposition of the reality of space itself.

So that if space is relational (i.e., not absolute), then position is relative—presumably this is because there simply are no parts of space for position to depend on. Thus the existence of absolute position implies that space is absolute.

In his reply, Clarke reiterates his argument as an attack on Leibniz's reading of the Principle of Sufficient Reason (p. 30); he then insists that "different places are really different or distinct from one another, though they be perfectly alike." I take this to be a tacit endorsement of the inference from absolute space to absolute position. He then goes on to attempt a feat of Leibnizean jujitsu on Leibniz's argument that relational space implies relative position:

> there is this evident absurdity in supposing space not to be real, but to be merely an order of bodies; that… if the earth sun and moon had been placed where the remotest stars now are… it would not only have been, (as this learned author rightly says) *la même chose*, the same thing in effect; which is very true: but it would also follow that they would have been in the same place too, as they are now: which is an express contradiction.

So it appears that Clarke, like Leibniz, endorses both conditionals: space is absolute if and only if position is absolute.

A similar series of exchanges takes place over the issue of absolute motion. Clarke realizes that on Leibniz's view, motion as well as position must be relative:

> If space was nothing but the order of things coexisting; it would follow that if God should remove the whole material world entire, with any swiftness whatsoever, yet it would always continue in the same place: and that nothing would receive any shock upon the most sudden stopping of that motion.[15] (p. 32)

Clarke seems to be trying to shift the arena of argument from metaphysics to physics: this dynamic shift is meant not as an attack on Leibniz's Principle of Sufficient Reason, but as a challenge to provide an adequate treatment of dynamics without the supposition of absolute space. Leibniz very wisely refuses to engage at this level, and retorts that the new argument,

---

[15] Although Clarke is usually given credit for having invented this 'dynamic shift' argument, its origin is actually much earlier. In 1277, under instructions from Pope John XXI, the bishop of Paris condemned 219 propositions, and declared that acceptance of any of these propositions was punishable by excommunication. Among the propositions condemned were: "That the only wise men of the world are philosophers"; "That by nutrition a man can become another numerically and individually"; and "That God could not move the heavens [that is, the world] with rectilinear motion; and the reason is that a vacuum would remain" (see Grant 1974, pp. 45-50). From the fourteenth century onwards, it was observed (by, e.g., Buridan and Oresme) that the denial of this last proposition implies the existence of absolute motion (see Grant 1979, especially pp. 230 and 243).

like the old one, conflicts with the Principle of Sufficient Reason, and thus only serves as a reductio of the chimerical supposition of absolute space (p. 38).

Clarke's strategy in replying is entrenchment:

> Two places, although exactly alike, are not the same place. Nor is the motion or rest of the universe the same state; any more than the motion or rest of a ship, is the same state, because a man shut up in the cabin cannot perceive whether the ship sails or not, so long as it moves uniformly. The motion of the ship, though the man perceives it not, is a real and different state, and has real different effects; and, upon a sudden stop, would have other real effects; and so likewise would an indiscernible motion of the universe. (pp. 48-49)

He supports this by appealing to Newton, and his program in the *Principia* of finding methods for determining the true motions of bodies. His concludes: "The reality of space is not a supposition, but is proved by the foregoing arguments…." This is going too far, since Leibniz is not likely to agree that Newton's physics is the only way to save the phenomena. But it is at least clear that Clarke thinks that motion is absolute if and only if space is absolute.

In his final letter (he died two weeks after writing it), Leibniz adds little new: Clarke's dynamic shift leads to a violation of the Principle of Sufficient Reason, thus it follows that his supposition of absolute space must be false. But he does give a very interesting reply to Clarke's comments about the man in the ship:

> The author replies now, that the reality of motion does not depend upon being observed; and that a ship may go forward, and yet a man, who is in the ship, may not perceive it. I answer, motion does not indeed depend on being observed; but it does depend on being possible to be observed. There is no motion, when there is no change that can be observed. And when there is no change that can be observed, there is no change at all. (p. 74)

Unfortunately, Leibniz is less explicit about the relationship between the reality of absolute space and the absoluteness of motion than he was about the absoluteness of position. But he is clear that Clarke's story about the dynamic shift works only if the existence of absolute space is granted. So it seems pretty certain that Clarke and Leibniz agree that motion is absolute if and only if space is absolute. This is a pleasing result—and a surprising one, given the obscurity of their respective views on the existence of space.

(iii) Locke.

There are modern-day substantivalists—see Butterfield 1989b, Hoefer 1996, and Maidens 1993—who believe in substantival spacetime, but follow Leibniz in claiming that if all material bodies were shifted three metres to the left, it would be *literally* la même chose, and not merely the same thing in effect. This position is clearly motivated by the desire to avoid the anti-substantivalist conclusion of the hole argument. So it is tempting to dismiss these authors as revisionists. To counter this temptation it is helpful to have in mind the views of Locke, who appears to have been a substantivalist while denying the possibility of absolute motion. Thus, the connection that Leibniz and Clarke found between absolute space on the one hand and absolute position and motion on the other was by no means obvious or inescapable to their contemporaries.

Space is a difficult topic for Locke. In his *An Essay Concerning Human Understanding*, Locke mentions several times that there is something very abstruse and peculiar about it (Locke 1975, §§II.12.8, II.14.2, II.15.1, and II.15.9.fn). One point, however, is fairly clear: Locke's position on space is more like that of Newton and Clarke than that of Leibniz. In fact, Locke comes much closer to saying that space is a substance than do either of these so-called substantivalists.[16] In §II.13.17, he explicitly allows that space may be a substance:

> If it be demanded (as usually it is) whether this *Space* void of *Body* be *Substance* or *Accident*, I shall readily answer, I know not: nor shall be ashamed to own my Ignorance, till they that ask, shew me a clear distinct *Idea* of *Substance*.[17]

It becomes clear in the next paragraph that this answer is not entirely sarcastic. Locke argues that 'substance' must be used in three different senses in talking about his three

[16] Leibniz would disagree with this claim. The passage quoted above from the *New Essays* is Leibniz's commentary on this part of Locke's *Essay*, and it is clear that he feels that Locke requires only minimal correction at this point. I claim that Leibniz picks up on a passage in §II.13.26 where Locke leaves open the possibility of a relationalist account of space, and ignores the many passages that appear to flatly rule out such an account. The passage in question picks up on some themes which were developed in Locke's journal in 1687 and 1688, but which were largely abandoned in the Essay; see §8.4.k of Grant 1981.

I would like to note in passing that I have not been able to identify any seventeenth century thinker who was enthusiastic about the proposition that space is a substance.

[17] who are liable to make this demand of Locke; Bayle 1965, p. 382. This question was, however, already common in the sixteenth century; see §8.4.a of Grant 1981.

official categories of substance (God, finite intelligences, and bodies). He concludes: “And if they can thus make three distinct *Ideas* of *Substance*, what hinders, why another may not make a fourth?” I think that the message here is straightforward. Locke prefers to avoid the substance-accident game altogether—the following two paragraphs are a polemic concerning its futility. But if he is forced to play, then he will pronounce space to be a substance.

Locke goes on to speak of the “permanent parts of space” (§II.14.1; see also §II.15.1,2) and of these parts as being “at perpetual rest, one amongst another” (§II.13.14). At this point is seems that Locke is well-positioned to achieve his professed goal of being an “Under-Labourer” for the “incomparable Mr. Newton” (p. 10)—he has absolute space, and he should go on to endorse absolute motion.[18]

He instead introduces relative notions of place and distance: “in our *Idea* of *Place*, we consider the relation of Distance betwixt any thing, and two or more Points, which are considered, as keeping the same distance one with another, and so considered as at rest…” (§II.13.7). This allows him to define rest and motion with respect to material reference points: “when we find any thing at the same distance now, which it was Yesterday from any two or more points, which have not since changed their distance one with another, and with which we then compared it, we say it hath kept the same *Place*: But if it hath sensibly altered its distance with either of those Points, we say it hath changed its Place….”

Locke shows no interest in using his permanent parts of space to underwrite absolute position or motion. In fact, he devotes an entire paragraph to arguing that absolute position is nonsense:

> That our *Idea* of Place, is nothing else, but such a relative Position of any thing, as I have before mentioned, I think, is plain, and will be easily admitted, when we consider, that we can have no *Idea* of the place of the Universe, though we can of all the parts of it; because beyond that, we have not *Idea* of any fixed, distinct, particular Beings, in reference to which, we can imagine it to have any relation of distance;

[18] Mr. Newton seems not to have been entirely pleased with his Under-Labourer. Princess Caroline told Leibniz that “Neither Dr. Clarke nor Mr. Newton wishes to be thought a follower of Mr. Locke, but I cannot and would not be one of theirs”; Alexander 1956, p. 190. See Stein 1990 for some reasons.

but all beyond it is one uniform Space or Expansion, wherein the mind finds no variety, no marks. (§II.13.10)

He adds a further argument: "when one can find out, and frame in his Mind clearly and distinctly the Place of the Universe, he will be able to tell us, whether it moves or stands still in the undistinguishable *Inane* of infinite Space…." That is, absolute position implies absolute motion, which is absurd; therefore, no absolute position.

Locke believes in absolute space but not absolute motion. It seems, therefore, that he disagrees with the Leibniz-Clarke criterion for the absoluteness of space. What should we make of this? Probably just that Locke had not thought very deeply about these questions. But it is worth entertaining other explanations. For instance, there is a streak of verificationism in the passage quoted above from (II.13.10), which is echoed elsewhere in Locke's discussion of these issues:

Where and *When* are Questions belonging to all finite Existences, and are by us always reckoned from some known Parts of this sensible World, and from certain epochs marked out to us by the Motions observable in it. Without some fixed Parts or Periods, the Order of things would be lost, to our finite Understandings, in the boundless invariable Oceans of Duration and Expansion; which comprehend all finite Beings, and in their full Extent, belong only to the Deity. And therefore we are not to wonder that we comprehend them not…. (§II.15.8)

All of this is reminiscent of Leibniz's claim that "There is no motion, when there is no change that can be observed." One might be able to give a rational reconstruction of Locke according to which this sort of verificationism rules out the possibility of absolute motion, while still allowing our finite understandings to individuate the immovable parts of space.

3. Where does this historical survey leave us? Setting aside Locke's strange synthesis of absolute space and relative motion, it seems that we should identify substantivalism with the view that space has parts which endure through time, maintaining their identity. This means that substantivalism underwrites absolute notions of position and motion. In particular, substantivalists will endorse Leibniz-Clarke counterfactuals like "If all bodies were shifted three metres to the left it would be a different world." Relationalists, on the other hand, while acknowledging that space itself exists, deny that there are genidentical parts of space. They will, therefore, hold that Leibniz-Clarke counterfactuals are either false or meaningless, and deny that the nature of space underwrites any notions of absolute position or motion.

It is relatively easy to make sense of this sort of substantivalism about space. Substantivalists see space as enjoying a sort of existence analogous to that of material bodies, in that it is made up of existent parts which maintain their identity over time.[19] They also tend to think of spatial relations between bodies as being parasitic on the spatial relations between parts of space, in the manner that McGuire attributes to Newton: "in Newton's view, 'distance' is not merely a two-place predicate which characterizes a certain disposition among material objects. Rather, the distance (at a given time) between two objects, is the distance between those parts or regions in space which the objects in fact *occupy* at that time".[20] Furthermore, the physical role of substantival space is straightforward. For substantivalists, inertial effects are the result of absolute motion, which is in turn glossed as being motion relative to space. Thus they are able to explicate one of the central concepts of dynamics in a purely kinematic vocabulary.

Relationalism about space is slightly more difficult to assimilate. Sklar's gloss may be helpful here: "The claim is … that, properly understood, all spatial and temporal assertions should be seen not as attributing features to space or spacetime, but rather as attributing some spatial, temporal, or spatiotemporal relations to material objects" (Sklar 1974, p. 167). If you and I are three metres apart, it is because we are *three metres apart*, and not because we occupy parts of space which are three metres apart. Distance relations are primitive. And space is understood as the structure of possible spatial relations between bodies.[21] For traditional relationalists, of course, space was Euclidean, which is just to say

[19] Ordinarily, I will assume that these 'parts' are in fact points, so that substantivalism about space becomes the doctrine that the points of space exist and maintain their identity over time.
[20] McGuire 1982, p. 149. This attribution is presumably motivated by Newton's avowal in *De Gravitatione* that "The positions, distances, and local motions of bodies are referred to the parts of space"; Hall and Hall 1962, p. 137.
[21] Recently, relationalists have also been empiricists, so that relationalism and possibilia have been uneasy partners (see §6.12 of Earman 1989). As will become clear below, I am

that they thought that the distance relations between bodies were constrained to obey the rules of Euclidean geometry (see §3 of Barbour 1982 for a nice discussion of how access to the distance relations as data might lead one to postulate the existence of space). From this point of view, it is incoherent to imagine two possible worlds in which the relations between bodies are the same, but in which my location in one world is three metres to the left of my location in the other world. So relationalists must deny the truth of the Leibniz-Clarke counterfactuals. Furthermore, of course, it makes no sense to say that I am *now* three metres away from where I was an hour ago—so there is no motion relative to space in a relationalist world.

The remainder of this Chapter is given over to an evaluation of the relative merits of these two doctrines. The polemical situation is as follows. Prior to Stein 1967, it was commonplace to denigrate Newton for having posited the existence of a substantival space. The postulation of this "mystical philosophical superstructure" (Reichenbach's infamous phrase) was regarded as proof that Newton was at best hopelessly naive—and at worst relentlessly dogmatic—when he turned his attention away from physics in order to dabble in philosophy.[22] According to this view, Leibniz and Mach were long-suffering heroes who had to wait until the advent of relativistic physics in order for their insights into the nature of space and time to be fully appreciated.

This judgment has now been overturned. According to the current consensus, Leibniz's arguments against the existence of substantival space have force only if one accepts some version of the Principle of Sufficient Reason or the Principle of the Identity of Indiscernibles. These principles may be attractive, but they are metaphysical. So they can

quite happy to resolve this tension by cutting relationalism free from empiricism. Indeed, although I do not take any stand on the nature of modality, I believe that relationalism is compatible with any approach to this question—including full-blown modal realism. On this point, I am in full agreement with Lewis, who sees his brand of modal realism as being neutral between substantivalism and relationalism (Lewis 1986, p. 76 fn. 55). This neutrality will be exploited in the Appendix.

[22] The canonical presentation of this view is Reichenbach 1959; see Stein 1967, Stein 1993, and Earman 1970b for references to other standard statements. Cassirer was the most important dissenting voice among Reichenbach's contemporaries; see pp. 125-26 of Cassirer 1956.

never convince us to reject substantival space if it turns out that there are physical grounds for preferring it to relational space: “Unlike Leibniz, most of us do not have accepted metaphysical principles that we take more seriously than physics and from which we can hope to derive knowledge of space and time” (DiSalle 1994, p. 285). The epistemological arguments of Mach suffer a similar fate: ceteris paribus, they would be reasons to prefer a relational account, but physical arguments trump all others.

And, of course, the history of physics shows that substantival space is adequate in the non-relativistic domain, so that the burden of proof lies squarely with the relationalists. There is a wide variety of arguments which are meant to demonstrate the lack of cogency and/or inadequacy of relationalism. These are comprehensively surveyed in Earman 1989. Earman finds that the substantivalist “can point to three reasons for accepting a substratum of space-time points: the need to support the structures that define absolute motion, the need to support fields, and the need to ground the right/left distinction” (p. 173).

After considering and dismissing a couple of minor objections to the cogency of relationalism in the next section, I will turn to the arguments from the need for absolute acceleration and the need for fields. I will not consider the problem of incongruent counterparts and the foundation of the right/left distinction, since I am in essential agreement with Earman’s conclusion that this issue provides, at best, only very weak support for substantivalism (see Earman 1989, §7.5). The prima facie problem posed by fields will be described in §5; I will argue in the following section, however, that this problem is not genuine. The balance of the chapter will be taken up with the problem of absolute acceleration. Here, again, my conclusion will be that relationalism is not in any serious danger.

4. Leibniz believes that space exists, and that it has some determinate structure. But what does this latter claim mean? Well , “space denotes, in terms of possibility, an order of things which exist at the same time” (Alexander 1956, p. 26). So, since Leibniz is a plenist, it seems unproblematic for him to claim that space is Euclidean. That is, he can simply say: "At any given instant, if you look at the set of spatial relations satisfied by all pairs of point particles, you will find that they are isomorphic to the set of relations satisfied by all pairs of points in Euclidean space."

What account of the structure of space can be given by a relationalist who is not also a plenist? Suppose, for instance, that I have an ontology consisting of a finite number of point particles, which spend all of their time within some fixed distance of each other. The particles will only ever instantiate a minute fraction of the relations exhibited by the points of Euclidean space. So what sense does it make for me to claim the space is Euclidean?[23] At this point, I am liable to put quite a bit of weight on Leibniz's talk of *possibility*, and to say that space is the set of possible spatial relations between bodies ("space is nothing else, but that order or relation; and is nothing at all without bodies but the possibility of placing them"; Alexander 1956, p. 26). Manders 1982 provides the most plausible way of cashing this out, when he shows how a language quantifying over the points of Euclidean geometry can be replaced by one quantifying over possible finite configurations of point particles. In this way, any sentence about the relations between points of space that the substantivalist holds dear can be translated into a sentence about possible configurations of objects (see Mundy 1983 for a closely related alternative approach).

This approach has been criticized on the grounds that

> the empiricist tradition, from which most relationalists come, holds that laws of nature, dispositions, and potentialities all supervene on the actual, occurent facts. If correct, this means that there is no difference between a conservative relationalist, who holds that all there is to the world are such facts as that a body b_1 is five meters from body b_2 at time t, and a liberal relationalist, who holds that there are also subjunctive facts about what would happen if, say, b_3 were introduced between b_1 and b_2. (Earman 1989, pp. 135-36)

And, of course, the conservative relationalist is not entitled to assign any definite structure to infinite space if there is some upper bound on the distances between material particles. In the worst case, the relationalist would be unable to speak of empty space as having any determinate structure whatsoever.

[23] See §8.6 of Earman 1989 for a version of this objection.

Sklar considers this objection, and concludes that the best that the relationalist can do is to insist that

> his subjunctive talk is grounded not on the belief in an underlying space, but rather on the general truths about the actual relations he has inferred from experience. Since, he says, the structure of the actual relations among actual objects obeys say, Euclidean geometry, I can make assumptions about what would be the case if objects were where they are not…. (Sklar 1974, p. 172)

If there are no objects, however, Sklar thinks that the best the relationalist can do is to "simply deny the intelligibility of talking about totally empty space" (pp. 172-73).

I believe that the relationalist can do better than this, by replying that what these arguments reveal is not the constraints which relationalism places on possibility, but rather the poverty of this sort of simple-minded empiricism. Earman and Sklar seem to agree that the relationalist cannot entertain the existence of various possible worlds, all of which are empty but which have distinct spatial geometries. This claim is supported by arguing that if the relationalist could do this, then there would be dispositional properties which do not supervene on occurent properties, and this should be unacceptable to any good empiricist. But, of course, there is no necessary connection between relationalism and this sort of empiricism, any more than there is such a connection between substantivalism and realism (see §1.3).

Indeed, the claim that modal facts supervene on occurent facts is highly contentious. Consider the following story. Let's suppose that theoretical investigation reveals that the most natural framework for understanding the nature of the four fundamental sorts of physical interaction suggests the possibility of the existence of a *fifth* fundamental force. It appears that this mysterious fifth sort of interaction has never actually occurred in our universe, but that it *is* possible to "turn it on." The necessary apparatus is built, but just before the switch is flipped, it is realized that there are two possible ways the fifth force could interact with the other four: if the coupling constant is *e*, then everything will go well and turning on the machine will make Planck scale physics accessible to experiment; if the coupling constant is π, flipping the switch will result in a massive explosion which will destroy the solar system. Since no-one has any grounds for believing that the constant is *e* rather than π, the machine is dismantled without ever being activated, and the fifth force is never turned on.

Advocates of the supervenience of the modal on the occurent are committed to the view that there is no fact of the matter as to what would have happened had the machine been switched on—since to say otherwise is to admit that there are two possible worlds which agree on all of the occurent facts (the facts about the four interactions which are instantiated) but disagree about a modal fact (the value of the coupling constant). This strikes me as being quite counter-intuitive—I would like to say that there is a fact of the matter about whether the world would have been destroyed or not. Thus, I claim that supervenience must give out sooner or later, at which point we are driven to accept some primitive modal facts. This is by no means an idiosyncratic position—it is more or less required of anyone who rejects a strictly empiricist analysis of laws of nature (see Chapter 3 of Carrol 1994 for a spirited attack on supervenience which is driven by examples like the one used above).

There is no reason why contemporary relationalists cannot reject supervenience, despite the empiricist leanings of past relationalists. The relationalist, in accepting that there can be more than one kind of empty space, which differ from one another in no respect other than in what would be the case if matter were present, wants to allow supervenience to give out earlier than the substantivalist. To some, this may appear to be an advantage for the substantivalists—aren't they able to *explain*, where the relationalists are not, why this empty space is Euclidean and that one is hyperbolic? This is an instance of a type of situation which we will run across a number of times in what follows, where the substantivalists claim victory on the grounds of superior explanatory power.

The rhetorical structure here is for the relationalists and the substantivalists to agree that a certain relational structure holds, but where the relationalist is content to accept this as a primitive fact, the substantivalist claims to explain it in terms of the properties of substantival space. Of course, what the substantivalist sees as a genuine gain in explanatory power, the relationalist sees as being a cheap trick—by trading ideology for ontology, the substantivalist is able to 'explain' the relationalist's primitives.

I do not see that the substantivalists have any great advantage over the relationalists on the issue of the structure of unoccupied space—their 'explanation' strikes me as being no explanation at all. Similarly, it has been claimed that the relationalists whose ontology consists of point particles are at a disadvantage because they cannot explain why the triangle inequality holds for distances between particles in Euclidean space (see pp. 194-96 of Maudlin 1993 and §§4-10 of Field 1985; see also Bricker 1993). The idea here is that the relationalist must posit the triangle inequality as a primitive fact, whereas the substantivalist can stipulate that the distance between two points is the length of the shortest continuous path between these points, and then point out that the triangle inequality holds in virtue of the fact that the shortest path between A and C is no longer than the path formed by concatenating the shortest path between A and B with the shortest path between B and C. Again, this strikes me as being no explanation at all, and so I do not see that these considerations give any advantage whatsoever to the substantivalists. But, of course, Maudlin and Field *do* see an advantage here, just as Earman and Sklar see one arising out of the problem of empty space.

I do not believe that there is any principled way to resolve such disputes. The fundamental question is whether a certain metaphysical hypothesis should be accepted given its explanatory power. And here everything depends on intuitions. All I can say is that, personally, I do not see the problem of empty space, or the challenge to explain the triangle inequality, as detracting from the plausibility or cogency of relationalism.

5. We now turn to the first putative ground for the superiority of substantivalism over relationalism: that latter is inadequate because it is incompatible with classical field theory.

Formally, one represents a field as a function on space or spacetime (the function will be scalar-, tensor-, or operator-valued, depending on the nature of the field). Typically, one speaks as if the field were an assignment of properties to spacetime points: "A field in mathematical physics is generally taken to be a region of space in which each point … is characterized by some quantity or quantities which are functions of the space coordinates and of time" (Hesse 1961, p. 192); "The idea of a field enters as the idea that values of physical quantities can be attributed to the space-time points" (Teller 1995, p. 53). Note that the substance-accident dichotomy is being employed in a relatively naive manner here. I will return to this point in §6 below.

Naive or not, this way of speaking suggests that the relationalist is not entitled to field-theoretic physics. The point is put most sharply by Field, who asserts that:

> The problem for relationalism is *especially* acute in the context of theories that take the notion of *field* seriously, e.g., classical electromagnetic theory. From the platonistic point of view, a field is usually described as an assignment of some property, or some number or vector or tensor, to each point of space-time; obviously this assumes that there are space-time points, so a relationalist is either going to have to avoid postulating fields (a hard road to take in modern physics, I believe) or else come up with some very different way of describing them. (Field 1980, p. 35; see also Field 1985, p. 40).

Earman endorses this view, claiming that "In modern, pure field-theoretic physics, [the spacetime manifold] functions as the basic substance, that is, the basic object of predication" (Earman 1989, p. 155); he goes on to assert that relationalists must reformulate field theories in a "relationally pure" vocabulary in order to be entitled to talk about fields (p. 159).

The sort of reformulation that Field and Earman have in mind is in terms of action at a distance (this is very explicit in Field 1985). The paradigm here is Newtonian gravitational theory. Although this theory can be given the appearance of being a field theory if formulated in terms of the gravitational potential and Poisson's equation, it can also be viewed as an action at a distance theory. In the latter case, one need mention only the properties of actual material bodies, so that the theory is manifestly acceptable to a relationalist. Then the potential's prima facie assignment of properties to unoccupied spacetime points can be regarded as an innocuous fiction. Field's idea, which Earman seems to endorse, is that relationalists are entitled to employ only those theories which can be given an action at a distance formulation.

All of this looks very reasonable: if we have a theory which cannot be given a formulation in which it does not mention the properties of spacetime points, then it seems that this theory should be unacceptable to an orthodox relationalist.

There are a couple of ways that the relationalist can attempt to avoid this conclusion. According to the first, one begins by endorsing the view that the essence of field theory is the assignment of *dispositional* properties to parts of spacetime (see Stein 1970 for this reading of classical field theory). One then goes on to insist that there is no reason why relationalists should not employ such properties in making sense of the gravitational field, since they are already entitled to employ them in making sense of the structure of empty space. Thus the gravitational potential would just tell us what would happen to a material body if it bore certain relations to given existent bodies.

This strikes me as being a promising line of thought. It is not clear, however, that this ploy will work in all cases in which physicists would like to engage in field-talk. Does it make sense, for instance, for relationalists to speak of a non-trivial configuration of the electromagnetic field in matter-free Euclidean space? My worry here is that in order to speak of an inhomogeneous field, we need to be able to refer to distinct parts of space. But it seems that the relationalist may be committed to the idea that, in the absence of matter, there is no way to refer to the distinct parts of a homogeneous space. Rather than grappling with this problem, I will defend a more radical relationalist response in the next section.

6. These concerns should motivate us to look for a more satisfactory relationalist response (for similar approaches, see Malament 1982, fn. 11; Teller 1987, pp. 430-31 and Teller 1991, p. 382; and perhaps even §5.8 of Earman 1995). Here Field's idea—that there is a relevant difference between field theories which can be reformulated as action at a distance theories and those which cannot be so formulated—is very suggestive. Even by Field's lights, the relationalists are entitled to action at a distance theories. So we need only worry about parts of physics which are essentially field-theoretic.

In order to understand the nature of essentially field-theoretic theories, it is helpful to recall how fields found their way into physics (here I follow the standard philosophical treatments of the field-concept, Hesse 1961 and Nersessian 1984). Maxwell's theory of electromagnetism was the first true field theory—previously there had just been field-theoretic formulations of action at a distance theories (Newtonian gravitation); field theories that were regarded as merely phenomenological (fluid mechanics); and suggestive field-theoretic speculations concerning the nature of electricity and magnetism (Faraday). Maxwell and his immediate successors were immensely impressed by the fact that electrical effects are not instantaneous. If I wiggle my charge, it will cause your charge to wiggle, but not instantaneously. How then can energy be conserved? Where is the energy that will move your charge, once I have stopped moving mine? The most straightforward way to preserve the conservation of energy is to say that during the intervening time the energy is propagating through the ether. And this is just to say that the ether, in addition for providing a medium for the propagation of electric and magnetic waves, can store and release energy. It was thus natural for the Maxwellians to regard the ether as an infinite material body of unknown micro-constitution, which could be in a number of mechanical states. In particular, each region of the ether was supposed to possess kinetic and potential energy, presumably because of its elastic nature. So, on Maxwell's theory, a configuration of the electromagnetic field corresponded to a mechanical state of the ether.

This view was soon undermined. In the first instance, by Lorentz, who showed that Maxwell's equations were inconsistent with the notion that the ether was an ordinary elastic medium. Lorentz's ether was an infinite extended object, but could not be regarded as a mechanical system since it acted upon, but did not react to, ordinary matter. Nonetheless, he was wedded to the idea that the electromagnetic field was a state of a material ether. The ether was finally banished by Einstein, who showed that physics had no need for a preferred ether frame. Thus, the electromagnetic field could no longer be regarded as a state of some material medium.

At this juncture, we have two choices. We can accept Field's view, and regard the field as a property of spacetime, in which case relationalism appears to be in trouble. Or we can regard the field as an object in its own right, whose existence is on a par with that of material objects. The justification for this latter position being that since fields are spatially and temporally extended, as well as absorbing, storing, and releasing energy, they are quite similar to ordinary objects.

There is an immediate objection to this position. It is sometimes maintained that the essence of the transition to field theory lies in the rejection of the notion of genidentical parts (see, e.g., p. 178 of Cassirer 1956). It makes no sense, for instance, to ask whether this bit of field is the same bit that was hereabouts yesterday. Thus, since fields are crucially disanalagous to the paradigm case of substance (material bodies), they should be treated as predicates rather than subjects.

How are we to settle this question? Unfortunately, there is no precise and widely accepted explication of the substance-accident dichotomy. The parties to the substantival-relational debate use the terms in an informal manner. They have in mind, not the relatively precise Scholastic distinction, but the result of filtering this distinction through the mechanical philosophy of the seventeenth century and then letting it lie idle for several centuries. In particular, they have a picture of the material bodies of common sense as being the paradigm instances of substance, but are relatively vague about what other sorts of things might count as substances, and about the demarcation between substances and properties. In this context, it seems that we should be granted considerable freedom in handling things like fields, whose existence was undreampt of when the subject-predicate distinction first found its way into philosophy. Essentially, we have to decide which is more important: the points of resemblance between fields and material objects (spatial and temporal extension, causal efficacy) or the disanalogy concerning genidentical parts. And here we would do well to note that if we take the possession of genidentical parts to be a necessary condition for being a material substance, then we are committed to the claim that there are no material substances in a quantum world (see, e.g., Redhead 1990 and Teller 1995 for a discussion of the difficulties involved in individuating quantum particles).[24]

Thus, I believe that it is perfectly cogent for relationalists to claim that a field is a thing. And, of course, having done so, they are every bit as entitled to take the field, as they are to take material objects, to have parts which stand in certain determinate spatial relations to one another. And thus it makes perfect sense for them to speak of there being a certain sort of convoluted electromagnetic field located in a matter free Euclidean space, which has this sort of configuration at one point, and that sort at a point three metres off to the left. Indeed, it seems clear that once the field is admitted into the relationalist ontology there cannot be any particular problems raised by field-theoretic physics, since the field is just a special sort of extended material object—if relationalism was cogent in the first place, then it remains cogent in the presence of fields.

As far as I can tell, there is no compelling substantivalist counter-attack. Field has a very weak response to this relationalist ploy: "I have no doubt that dispensing with space-time in favor of 'parts of a field' is possible, so if one wants to call this a version of relationalism then relationalism is certainly a manageable position. … It seems to me however that this 'saves' relationalism only by trivializing it" (Field 1985, p. 41). But on closer examination, Field has no grounds for regarding this as a trivial form of relationalism, other than a reiteration of his claims that a field is not an object but an assignment of properties, and that relationalist physics is actions at a distance physics (pp. 41-42, especially fn. 21). Perhaps what he has in mind here is the claim that if relationalists are allowed to help themselves to field-talk, then there will no longer be any interesting differences between relationalists and substantivalists. If this were true, then it might well be

[24] Here, strictly speaking, I here depart from my policy of relying on intratheoretic considerations in looking for an interpretation of Newtonian physics. There are two ways that this move can be scored. The first is as a genuine departure from this policy, and thus as an illustration of my major claim of Chapter 8—that intertheoretic considerations have important interpretative consequences. But it is also possible to count the discussion at hand as being intratheoretic in a broader sense—the issue at hand is the interpretation of Newtonian spacetime structures, and one of the theories set in such structures is many-particle quantum mechanics.

a sufficient reason for denying relationalists this freedom. But, of course, there remain many points of disagreement between the two factions, even if relationalists are granted the right to talk about fields. One of the most obvious is the central disagreement over the status of the Leibniz-Clarke counterfactuals. In Chapters 7 and 8, I will argue that this is also one of the most important points of disagreement, since it provides the best hope for re-introducing physical content into the substantival-relational debate.

The moral of the supposed problem of fields is not that relationalism is incompatible with field-theoretic physics. It is, rather, that there is a danger in attempting to read off interpretation from formulation. The 'problem of fields' only looks interesting because people take too seriously the fact that fields are *represented* by assigning properties to the points of the mathematical object which represents spacetime. But, as Malament points out, "one could characterize a sofa similarly" (Malament 1982, fn 11). That is, one could describe ordinary objects as being properties of the parts of spacetime that they occupy.[25] Here, as elsewhere, it is a mistake to move too quickly from formalism to interpretation.

7. We now turn to the second major challenge to the adequacy of relationalism about space: the problem of absolute acceleration. The best approach to this problem is through Stein 1967. Cast your mind back to those dark days when Newton was derided for having posited absolute space, and furthermore, was taken to be either hopelessly naive or inexcusably metaphysical in comparison with those heroes of empiricism, Leibniz and Mach. Stein provides a devastating refutation of this evaluation of the classical substantival-relational debate.

In modern idiom, the insights of Leibniz and Mach turn upon the suspicious character of the concept of absolute velocity. If absolute space exists, then it follows that every object has an absolute velocity—the velocity of the object relative to absolute space.

[25] Indeed, in *De Gravitatione*, Newton seriously entertains the view that material bodies are simply regions of space which God has chosen to endow with impenetrability; Hall and Hall 1962], pp. 139-40. See Stein 1970 for an enlightening discussion.

Yet absolute velocity plays no role in Newton's physics: you could boost the velocity of every object in the universe by some fixed quantity, and the resulting situation would be indistinguishable from the original.

This has often been taken to be proof that the postulation of absolute space was a mistake. Stein argues that although there is a sense in which this is true, it is a mistake that no one can blame Newton for having made. The argument goes like this. We can agree that, since absolute velocity does no work, there is no need for absolute space. But we are faced with the need to account for inertial effects, and the most straightforward method for doing so is to posit a notion of absolute acceleration. Newton believed that he needed absolute velocity in order to have absolute acceleration, and absolute space in order to have absolute velocity. So he thought that he needed absolute space in order to account for inertial effects (see §2.i above). On this account, far from being naive or dogmatic, Newton was a sophisticated natural philosopher. He posited an unobservable entity (absolute space) which would entitle him to the concept (absolute acceleration) that he needed in order to account for the phenomena (inertial effects).

Of course, from the four dimensional perspective, we can see that this is a mistake, since if we work with neo-Newtonian spacetime rather than Newtonian spacetime, we get exactly what we need—absolute acceleration without absolute velocity—so that we can explain inertial effects while doing away with absolute space.[26] Thus, with hindsight, we can

---

[26] See Chapter 2 of Earman 1989 for a precise characterization of various classical spacetimes. Informally, we may describe the main options as follows. The most basic is Leibnizean spacetime: the idea is that we have a set of spacetime points with the topological structure of $\mathbf{R}^4$; we assume that there is an absolute notion of simultaneity, so that our spacetime is foliated by hyperplanes of simultaneity in an unambiguous way; on each hyperplane of simultaneity, there is a spatial metric (giving the spatial distance between simultaneous spacetime points); and there is a globally defined temporal metric (giving the temporal interval between any two spacetime points). In Leibnizean spacetime, there are no notions of absolute acceleration or absolute velocity. Adding a flat connection/affine structure, which picks out the class of 'straight' timelike lines gives you neo-Newtonian spacetime. In neo-Newtonian spacetime, there is no notion of absolute velocity, but there is a notion of absolute acceleration (which measures how far a worldline is from being 'straight'). Finally, one can move all the way to Newtonian spacetime by defining a notion of distance between non-simultaneous points. This gives you both absolute acceleration and absolute velocity.

see that although there was a deep insight in the relationalist critique of absolute velocity, it was one which the relationalists themselves were in no position to exploit. The deep insight was that Newton had postulated more structure than he needed. But, according to Stein, the relationalist program of constructing a physical theory based on the spatial relations between bodies—rather than on the relations between bodies and substantival space—was doomed to failure. This is because it follows from fact that a notion of absolute acceleration is needed in order to account for inertial effects that "a 'connection' is indeed required for the formulation of the principles of dynamics, and that *it cannot*—as Newton quite clearly indicates—*be defined in terms simply of the spatial relations between bodies*" (Stein 1967, p. 187; emphasis in the original).

There are two ways in which relationalists can answer Stein's version of Newton's *De Gravitatione* argument: head-on by showing that their ontology of bodies and their ideology of spatial relation between bodies actually *do* contain the resources to account for inertial effects (Machianization); or obliquely, by showing how this can be achieved by maintaining their ontology while enlarging their ideology (Sklar's Maneuver). In the next two sections, I will argue that the latter is successful in rescuing relationalism from the problem of absolute acceleration. I will then turn to an evaluation of the program of Machianization. My conclusion is that it enjoys only partial success.

8. Sklar recognizes the challenge that absolute acceleration poses for relationalism, and explicitly acknowledges that Machianization is incapable of meeting this challenge (Sklar 1974, p. 229; and Sklar 1976, pp. 14-15).He goes on to advocate a type of relationalism about space which he believes can handle the existence of inertial effects (Sklar 1974, §III.F.2; see also §IV.D of Sklar 1976). The fundamental idea is that the property of absolute acceleration may be included in one's ideology without the foundation of an underlying ontology of substantival spacetime:

> To maintain the relationist doctrine of space and time in the face of the acceptance of absolute motions, what we must do is deny that the predicate "is absolutely accelerated" is a relational term! The expression

> "A is accelerated" is incomplete . To complete it, we must answer the question, "Relative to what is A accelerated?" But the expression "A is absolutely accelerated" is a complete assertion, as is, for example "A is red," or "A is bored," and unlike "A is north of." (p. 230)

Here spin might provide a better analogy than color. The Sklar-relationalist wants to insist that particles can have degrees of freedom which are "internal" in the sense that they are not to be thought of as kinematic degrees of freedom—although they are, as a result of an unfortunate historical accident, called "absolute acceleration." This may be strange, but it is not conceptually incoherent, since it is just what is ordinarily said about classical systems with spin. On this account, absolute acceleration is not to be thought of as acceleration relative to substantival space, and is thus no longer antithetical to relationalism.

Of course, one need not stop there: one can use the same trick to become a relationalist with a notion of absolute velocity (for this suggestion, see Horwich 1978, pp. 402-03; and Maudlin 1993, p. 187). In this case, the relationalist insists that absolute space is a chimera, but that objects have a primitive property called absolute velocity, which is not to be thought of as velocity relative to anything. Of course, the respective absolute velocities and accelerations of various objects satisfy some interesting relations—e.g., absolute accelerations are determined by Newton's law of gravitation, and if two objects have the same absolute acceleration and are at rest with respect to one another then they have identical absolute velocities.

At this point, it seems that we are once again landed in type of dilemma discussed in §4: both the relationalists and the substantivalists claim that they are entitled to a certain amount of ideology; substantivalists then go on to assert that the explanatory superiority of their ontology justifies them in disallowing the relationalists' claim.

In this case, I think that nearly everyone shares the intuition that the substantivalists are correct in claiming that they—and they alone—are entitled to a notion of absolute velocity. This is because it is natural to regard the substantivalist's kinematic explication of absolute motion as motion relative to space as a genuinely *physical* explanation. This is in sharp contrast to the explanations which were dismissed in §4 as being patently *meta*physical.

The case of the relationalist who accepts absolute accelerations but not absolute velocities is a bit more complicated. Here relationalists *do* have an advantage over substantivalists: they posit exactly as much dynamical ideology as necessary, without being committed to any excess ontology. What Sklar-relationalists are really doing is offering a non-standard interpretation of the neo-Newtonian formulation of non-relativistic physics. The standard interpretation of neo-Newtonian spacetime includes in its ontology a substantival spacetime (see Chapter 3). The Sklar-relationalists refuse to buy into this four-dimensional interpretation, offering in its stead an interpretation in terms of a relational space. Under this interpretation the affine connection of neo-Newtonian spacetime does not appear as a four-dimensional spatiotemporal geometric structure, but serves rather to underwrite the 'geometry' of the family of absolute acceleration properties.

Earman 1989 provides the only strong objection to this version of Sklar's maneuver:

> Of course, the [neo-]Newtonian apparatus can be used to make the predictions and afterwards discarded as a convenient fiction, but this ploy is hardly distinguishable from instrumentalism, which, taken to its logical conclusion, trivializes the absolute-relational debate, as it does so many issues in the philosophy of science. (p. 128)

He challenges the Sklar-relationalists to formulate a physical theory in a relationally acceptable vocabulary. That is, he demands that they reformulate Newton's theory so that it does not mention spacetime or substantival space (see especially §§6.9 and 8.7 of Earman 1989).

Strictly speaking, it is very unfair to accuse Sklar-relationalists of instrumentalism. After all, they accept the existence of absolute acceleration, a highly theoretical concept which is by no means immediately observable. This, of course, would be anathema for a true instrumentalist. I believe that Earman is motivated by the idea that the *natural* interpretation of Newtonian or neo-Newtonian spacetime is in terms of a substantival space or spacetime, where 'natural interpretation' means something like "what you get if you attempt to read ontology directly off of the mathematics."[27] And he believes that this means that relationalists are entitled to employ this mathematical apparatus only if they can show how it could in principle be replaced by a formulation of the theory whose natural interpretation is relationalist. His idea is that unless this can be done, the Sklar-relationalists appear to be underhanded, and perhaps even to have instrumentalist leanings, since they refuse to take seriously most of the theoretical structure of their science.

I am willing to grant the claim that the natural interpretations of Newtonian and neo-Newtonian theories are substantivalist—if we try to read our ontology off from our mathematics, then we will believe in spacetime points. But I would insist that this is not to say that there are no other viable interpretations. And there is no reason, so far as I can see, to believe that an interpretation of a given formulation of a theory is viable only if it is the natural interpretation of some formulation. Thus I am quite content to believe that Sklar-relationalism is a legitimate interpretation of Newtonian gravitational theory formulated in neo-Newtonian spacetime, even though there is no known way to bypass the mention of a spacetime.

9. It seems, then, that we have two viable three dimensional interpretations of non-relativistic physics: the standard substantivalist interpretation of theories formulated in Newtonian spacetime; and the Sklar-relationalist interpretation of theories formulated in neo-Newtonian spacetime. I do not believe that the problem of fields points up any reason for preferring one of these interpretations over the other.

The problem of absolute acceleration, on the other hand, does motivate two important objections—one which bears against substantivalism, and one which bears against relationalism. The substantivalist, of course, stands accused of postulating excess ideology—physics has no need of absolute velocity. The relationalists, on the other hand,

[27] Of course, this sense of 'natural' is highly conditioned by the present state of philosophy of space and time.

don't seem to postulate enough ontology, since it would seem that they cannot *explain* the notion of absolute acceleration in the neat manner that the substantivalists can.

There are a number of other objections that can be brought to bear against each interpretation. One can add to the metaphysical objections mentioned in §4, the accusation that relationalism presupposes a more robust realism about mathematical entities than substantivalism does (Field 1985). On the other hand, as discussed in the Leibniz-Clarke Correspondence, substantivalists must violate the Principle of the Identity of Indiscernibles in postulating an infinite number of distinct but indistinguishable states of affairs.

At this point, I would like to declare the match a draw. We have pursued the connections between physics and the substantival-relational debate and found that, modulo questions of metaphysics, both positions are viable as doctrines about space in the non-relativistic context. Rather than pausing here and examining the purely metaphysical merits of the two positions, I would like to press on. This is because: (i) the metaphysics is difficult; (ii) I do not believe that either side has a decisive advantage; (iii) I believe that the arguments from the non-relativistic context which have the most direct bearing on my primary goal are motivated by physics rather than metaphysics.

There is one bit of business remaining before we can bring to a close this examination of three dimensional interpretations of non-relativistic physics. At the end of §7, I mentioned that there are two strategies that relationalists can pursue in order to avoid the problem of absolute acceleration. The first is Sklar's Maneuver which, I have argued, is a success. The second is the program of Machianization, in which one tries to make do with traditional relationalist ontology and ideology without supplementing it with a primitive notion of absolute acceleration. The remainder of this chapter will be given over to an evaluation of this latter program. Ultimately, I will conclude that Machianization enjoys only partial success, and so will have little bearing on the outcome of the non-relativistic substantival-relational debate. It is nonetheless worthwhile to undertake a detailed evaluation of the program, because this will lead us to develop a framework for discussing relationalist formulations of physical theories which will be helpful in Part II.

10. Our starting point is Stein's comment to the effect that a connection cannot be constructed on the basis of the relative distances between particles. My first step is to develop the standard framework in which this issue is discussed in the philosophical literature. In the next section, I will argue that this framework is inadequate. I will then go on to summarize the more adequate treatment which this issue has been given in the physics literature. This will lead to the conclusion that there is no Machianization of the full Newtonian theory.

It seems that the natural interpretation of a theory formulated in Leibnizean spacetime a relationalist one, just as the natural interpretation of a theory formulated in Newtonian spacetime is a substantivalist one. The idea is that we may regard Newtonian spacetime, with its preferred standard for identifying spatial points over time, as a natural representation of substantival space. We can, in fact, imagine Newtonian spacetime as being built out of a stack of instantaneous spaces, together with a method of identifying spatial points over time. In the same way, we may regard Leibnizean spacetime as being comprised of a set of instantaneous spaces without any standard for identifying points on distinct simultaneity slices. From this perspective, Leibnizean spacetime is the natural representation of a relationalist world in which the parts of space do not even exist, let alone maintain their identity over time. Notice that neo-Newtonian spacetime does not fit into this scheme: there is no straightforward way to see its affine structure as arising out of the nature of the parts of its instantaneous spaces. Which is just to say that neo-Newtonian spacetime does not have a natural three dimensional interpretation.

Now it seems like we have a way of making Stein's remarks precise. Newton o lute acceleration in order to account for inertial effects. This seems to imply that we need a connection on our spacetime. But (non-Sklar-)

relationalists are committed to doing their physics in Leibnizean spacetime, which has no connection. So there cannot be an empirically adequate relationalist physics.

Unfortunately, this elegant argument is far too quick. Let's suppose that we accept the convention that the relationalists are committed to Leibnizean spacetime. This means that they should not posit additional spacetime geometry. But this does *not* mean that they cannot define a notion of absolute acceleration. There remains the possibility of attempting what Friedman calls 'Machianization':

> it may be possible to give a relationalist account of the affine properties of space-time: that is, a reduction of the affine properties of spacetime to some physical properties of the total distribution of matter. It may be possible thereby to exhibit the distorting effect experienced by an accelerating or rotating object as depending directly on the relations of that object to the rest of the matter in the universe, bypassing any appeal to a privileged class of … inertial trajectories. (Friedman 1983 , p. 67)

That is, relationalists may circumvent Stein's objection by postulating an affine structure in order to account for inertial effects—so long as this affine structure is definable in terms of the relative distances and velocities of material bodies rather than being part of the fixed spatiotemporal geometry.

11. The business of this section is to counter an argument that suggests that Machianization is fully successful, in the sense that we can find a formulation of Newtonian gravitational theory in which the affine structure depends on the matter distribution, and which admits of a relationalist interpretation. This will highlight a disadvantage of the convention of associating Leibnizean spacetime with relationalism, and will lead us to a sharper formulation of the program of Machianization.

If we are doing Newtonian gravitation via Poisson's equation in neo-Newtonian spacetime then the equivalence principles tells us that we have extraneous structure: since switching from one frame to a second accelerated (but non-rotating) frame merely transforms the description that observers give of the gravitational field, we find that there are a variety of pairs (flat affine connection, gravitational potential) which all give rise to the same set of spacetime trajectories of massive bodies. In this sense, the choice of a connection and a potential is conventional. This conventionality can be eliminated by absorbing the potential into the connection, while leaving the rest of the structure of neo-Newtonain spacetime untouched. This gives us the Newton-Cartan theory of gravity, in which massive bodies follow the geodesics of a non-flat affine connection (see §III.4 of Friedman 1983 for the details).

The spacetime of the Newton-Cartan theory consists of the fixed structure of Leibnizean spacetime together with a connection which depends on the matter distribution. This suggests that the Newton-Cartan theory is a Machian theory. We can reinforce this suggestion by introducing the notion of the symmetry group of a theory (Friedman 1983, §II.2). The idea is to associate with each classical spacetime theory the group of transformation of the theory's spacetimes which preserves all of the *fixed* spacetime structure. Thus the symmetry group of a theory formulated in Newtonian spacetime just consists of time-independent translations and rotations. In the case of the Newton-Cartan theory, the elements of the symmetry group are required to preserve the Leibnizean structure, but not the affine structure. This means that the rotations and translations which the elements of the group induce on each simultaneity slice can now be independent of one another. And Earman has given a general argument, along the lines of that of the previous section, which is supposed to show that any theory whose symmetry group is Leibnizean should be given a relationalist interpretation (see §III.3 of Earman 1986; also Earman 1989, §3.6 and Stein 1977, p. 8).

The problem is that the Newton-Cartan theory is *not* a properly Machian theory. This point is made by a simple thought experiment: imagine a possible world whose only contents are two massive globes and a cord connecting them, and in which there is initially no relative motion. If this world is described by ordinary Newtonian mechanics, then we know there are at least two things that could happen: the system could be non-rotating, in which case it will collapse, due to its own gravity; or it could be rotating fast enough that the cord stays taut, and there is never any relative motion. Since the Newton-Cartan theory is an empirically equivalent reformulation of Newtonian mechanics for finite matter distributions with appropriate boundary conditions, both of these options exist according to this supposedly Machian theory. But in a truly Machian theory, the initial relative motions must determine the relative motions for all time, so that there is only one possibility—presumably that the system should collapse.

Although the affine connection of a Newton-Cartan spacetime *depends* on the matter distribution, it is a mistake to think that it is *determined* by it. If the world of our thought experiment is described by the Newton-Cartan theory, then the globes help to determine the affine structure—but it is always possible to decompose this structure into a fixed background structure and a portion contributed by matter.[28] The notion of spacetime symmetry is inappropriate as a guide to interpretative questions because it is incapable of distinguishing between structure which is determined by the matter distribution, and structure which merely depends on the matter distribution.[29] Thus, the idea that a theory is Machian iff its fixed structure is Leibnizean is unacceptable.

12. All of this suggests that we need a sharper criterion for deciding when a theory is truly Machian. The most promising approach is that of Barbour and Bertotti 1982 (see also Barbour 1982, Barbour 1994a, and Lynden-Bell 1995).[30] Here the idea is to attend to the configuration space of the theory rather than to its spacetime structure. Start with an N particle configuration space, Q. Here a point in Q represents the positions of the N particles

[28] If there is an infinite amount of matter, then there are no non-conventional grounds for performing this decomposition, but this is not to say that it is impossible; see Malament 1995 and Norton 1995.

[29] I believe that there is a closely related group which *is* relevant to interpretative questions: the indistinguishability group; see §IV.5 of Friedman 1983, but notice that one needs to generalize Friedman's treatment in order to assign non-trivial indistinguishability groups to theories which utilize curved spacetimes. This group maps privileged frames onto privileged frames. In the context of the equivalence principle, the class of privileged frames will be the non-rotating frames. Thus the indistinguishability group of the Newton-Cartan theory is the so-called Maxwellian group (Earman 1989, §2.3), which is smaller than the Leibnizean group (it contains time-dependent translations but not rotations). See Kuchar 1980 and Christian 1994 for the role of this group in the physics of the Newton-Cartan theory.

[30] The following account glosses over some technicalities, but is not misleading in any important respect. See examples 5.5, 5.6, 6.2, and 6.3 below for the details.

relative to some fixed inertial frame. We now construct the phase space of our theory, T*Q, the cotangent bundle of Q. Once we write down a Hamiltonian for the theory, we can find the trajectories in T*Q which represent physically possible histories.

A good relationalist will object that each physically possible state of the system is represented by many points of Q. That is, for each set of relative distances between the points, there will be many points of Q which represent the particles as having those relative distances, but differ as to their location relative to the inertial frame. Similarly, there will be many curves in T*Q which represent a single sequence of relative distances and velocities of the particles.

A healthy relationalist reaction is to advocate replacing Q, the set of configurations relative to an inertial frame, with Q_0, the set of relative configurations. Each point in Q_0 represents the relative distances of the particles at a given time, rather than their positions relative to an inertial frame. Q_0 can be constructed by quotienting Q by the action of the group of orientation-preserving Euclidean isometries. The goal is then to formulate an empirically adequate relationalist dynamics by writing down a Hamiltonian on T^*Q_0 which determines a set of physically possible trajectories in the space of relative positions and velocities.

The discussion of the previous section shows that the Newton-Cartan theory is not a such a theory, since specifying an initial point in T^*Q_0 does not determine the subsequent evolution. In this way, we can see that Machianization must fail: no theory which is formulated on T^*Q_0 can be empirically equivalent to a Newtonian theory of gravitation, since no such theory can reproduce the Newtonian theory's prediction that there is more than one possible future for any given set of initial relative positions and velocities (a version of this observation can be found in §3.4 of Earman, and Friedman 1973).

But Stein's claim that no Machian theory can account for inertial effects does not follow. Barbour and Bertotti 1982 define a Hamiltonian on T^*Q_0 which yields dynamical trajectories which are empirically indistinguishable from those of the Newtonian theory iff the center of mass of the universe is non-rotating (this is particularly clear in the formulation of the theory given in Lynden-Bell 1995). Here, ‘empirical equivalence’ means that, given the same set or relative distances and velocities as initial data, the two theories predict identical relative distances and velocities. Thus, while examples such as the rotating globes show that there is no Machianization of Newtonian gravitational theory, the work of Barbour and Bertotti shows that there is a perfectly cogent relationalist interpretation of the fragment of Newtonian theory in which the center of mass is non-rotating.

Furthermore, the currently available evidence indicates that the angular momentum of the universe is vanishingly small (Lynden-Bell 1995, p. 176). So we have no empirical grounds on which to prefer Newton’s theory to that of Barbour and Bertotti. Thus, in our world, inertial effects may indeed be referred to the fixed stars rather than to absolute space. Of course, as Earman emphasizes, there is much more to classical physics than gravitation, and ultimately the relationalist program must be judged in terms of its success in accounting for all classical phenomena (Earman 1989, pp. 95-96). So far, however, the outlook is encouraging, since Barbour and Bertotti are able to formulate a relationalist theory of electromagnetism which is empirically equivalent to Maxwell’s theory in the non-rotating case (p. 303; see also pp. 261-62 of Smolin 1991).

What does this leave Machianization as a means of saving relationalism about space from the problem of absolute acceleration? If the goal of relationalism is to provide an interpretation of a sufficiently large fragment of Newtonian theory, or of a theory which is empirically adequate in the domain of Newtonian theory, then the work of Barbour, Bertotti, and others on relationalist dynamics vindicates relationalism about space. In this case, Machians emerge triumphant from the substantival-relational debate. This is a momentous conclusion, and one that has received far too little attention in the philosophical literature.

For my purposes, however, it is convenient to follow standard practice, and to fault Machianism on the grounds that it cannot provide an adequate interpretation of the full Newtonian theory. From this perspective, neither relationalists nor substantivalists have provided an adequate three dimensional interpretation of Newtonian physics. This motivates the four dimensional turn in the philosophy of space and time, which is the subject of the next Chapter.

CHAPTER 3
The Non-Relativistic Debate: Spacetime

1. The previous Chapter surveyed one half of the of the non-relativistic substantival-relational debate. I examined the viability of substantivalism and relationalism construed as doctrines about the nature of the existence of space. In this Chapter, I complete my discussion of the non-relativistic debate by considering four dimensional versions of these interpretative strategies. Again, I will be in search of an interpretation that is capable of handling all of Newtonian physics. My discussion will focus on point particle mechanics set in neo-Newtonian spacetime. This will allow me to draw out all of the issues which are required for the discussion of GR in the next Chapter.

It seems to be universally agreed in the philosophical literature that four-dimensional interpretations are preferable to three dimensional interpretations. This preference is justified by appeal to two distinct sorts of ground. The first is the feeling that the three dimensional doctrines are in conflict with the lessons of relativistic physics, since they presuppose the existence of a preferred foliation of spacetime by surfaces of simultaneity. In Chapter 8 we will see that this judgment, while natural, is premature, since quantum gravity may provide grounds for favoring three dimensional interpretations of GR. This issue can be set aside for the time being, however, since we are here interested only in considerations which are internal to Newtonian physics itself (see §1.4).

The second ground for preferring four dimensional interpretations emerges from the discussion of the previous chapter. Both substantivalism about space and relationalism about space are commonly believed to founder on the problem of absolute motion. This is most clear in case of three dimensional substantivalism. All hands recognize that a concept of absolute acceleration is required in order to do interesting physics. Their belief that the parts of space maintain their identity over time gives substantivalists about space an ontological underpinning for this crucial concept. But it also underwrites a scientifically otiose notion of absolute velocity. The move to substantivalism about spacetime is motivated by the hope that this will yield an interpretation of the structure of neo-Newtonian spacetime which legitimizes absolute acceleration but not absolute velocity, while somehow managing to remain faithful to the spirit of three dimensional substantivalism. The story about relationalism is slightly more complicated. As was shown in Chapter 2, it *is* possible to give a three dimensional relationalist interpretation of a theory set in neo-Newtonian spacetime, and thus to handle the problem of absolute acceleration without committing oneself to the existence of absolute velocity. But, it is often maintained, absolute acceleration is actually a four dimensional notion, so that even for relationalists, the most natural interpretation is four dimensional rather than three dimensional (see especially Teller 1991).

The chief goal of this chapter is to reach an understanding of the strengths and weaknesses of the four dimensional interpretations both in comparison to each other and in comparison with their three dimensional relatives (see §§7 and 8, respectively). The first task, though, is to attempt to provide an explication of the content of substantivalism and relationalism about spacetime.

2. Even in their original three dimensional incarnations, substantivalism and relationalism were quite abstruse doctrines. There however, the analogy with material substances gave us some purchase. Three dimensional substantivalists can give some content to their thesis that space enjoys a substance-like existence by claiming that the parts of space, like the parts of bodies, endure and maintain their identities over time. The structure of space can then be taken to be the structure of the unchanging relations between the parts of absolute space. Relationalists about space can then contrast their position with substantivalism by denying that there are genidentical parts of space, and insisting that the structure of space is simply the framework of possible spatial relations between bodies (with no ontological underpinning). Thus the three dimensional substantival-relational debate can be given some straightforward content. This basic difference of opinion over the nature of the existence of space has an immediate consequence: a difference of opinion over the truth values of the Leibniz-Clarke counterfactuals. Best of all: the debate is intimately related to questions of physics via the problem of absolute motion.

Unfortunately, the situation is considerably less clear in the four dimensional context. The fundamental problem is that it makes no sense for substantivalists to claim that spacetime enjoys an existence like that of a material substance, since spacetime is four dimensional while material substances are three dimensional. At best, they could attempt to translate the analogy into the four dimensional context by insisting that spacetime enjoys a type of existence similar to that enjoyed by the worldtube of a material object. This analogy is not very close, however, since one of the most salient features of worldtubes is that they are striated by the worldlines of the parts of the object in question. But spacetime substantivalists will surely *deny* that spacetime is striated by the worldlines of the parts of space, since admitting the existence of such a preferred means of identifying points over time is tantamount to admitting the existence of substantival space—which is precisely what substantivalists were seeking to avoid in shifting their attention to spacetime.

It seems, then, that substantivalists cannot merely carry over to the four dimensional setting the terms of the three dimensional substantival-relational debate . But, of course, one *does* want a characterization of substantivalism about spacetime which is clearly the proper successor to three dimensional substantivalism. There are three prima facie promising proposals:

(i) The substantivalist about space or spacetime holds that it is legitimate to quantify over the parts of space or spacetime, whether or not they are occupied by matter (see especially Friedman 1983 and Field 1985).

(ii) Spacetime substantivalism amounts to the claim that spacetime is substance-like in that it subsists, in the sense that it would exist even if matter did not (see Auyang 1995, p. 137; Maudlin 1993, p. 184; and §§5.12.1 and 5.12.2 of Norton 1992a). Thus one often sees the claim that the existence of vacuum solutions to Einstein's field equations somehow vindicates substantivalism concerning the spacetime of GR (see, e.g., Earman 1970a, p. 314; and Hoefer 1996, p. 12).

(iii) Spacetime substantivalism is the denial of the following form of relationalism: "Spatiotemporal relations among bodies and events are direct; that is, they are not parasitic on relations among a substratum of … spacetime points that underlie events" (Earman 1989, p. 12).

The first two proposals are, of course, considerably more precise than the third. Unfortunately, neither of them amounts to an acceptable explication of spacetime substantivalism since both considerably distort the content of the three dimensional substantival-relational debate.

The first proposal suffers from some well known problems. The question 'What exists?' is equated with 'What domains are being quantified over?' But the latter is well-posed only relative to a choice of formal language. Since scientific language is very far from being a formal language, one expects that in general the choice of a formal language will be substantive, in the sense that different choices will lead to different conclusions about what exists. This is indeed the case when one applies this approach to the substantival-relational debate: Field has argued very convincingly that one has a choice between a language which quantifies over spacetime points but not numbers, and one which quantifies over numbers but not spacetime points (Field 1985). Thus, if we adopt the first approach to defining substantivalism we become involved quite quickly in the sort of traditional metaphysical questions—such as the existence of abstract objects—which I am trying to avoid. Furthermore, this approach suggests that there is little at stake in the substantival-relational debate when matter forms a plenum—since in this case quantification over material points and quantification over spacetime points will lead to expressively equivalent languages (Friedman 1983, pp. 222-23). But, as Earman notes, although Leibniz was a plenist he most certainly *did* believe that there was an issue between himself and the substantivalists (see §6.6 of Earman 1989).

Similarly, it is far from clear that relationalists should be interested in denying that either space or spacetime subsists in the sense of proposal (ii). As discussed in §2.4, Leibniz was in no way interested in denying that space subsists—he merely wanted to deny that its existence was analogous to that of a material substance. I do not see that anything should change when relationalism becomes a doctrine about spacetime: the relationalist will want to grant existence to spacetime, and there is no reason why subsistence should not follow.

This leaves proposal (iii) as the best of a bad bunch. It represents a translation into the four dimensional context of one of the fundamental intuitions of three dimensional substantivalism: that the ideology of spatiotemporal relations is underwritten by ontology. Unfortunately, where the ontological underpinning of spatial relations could be compared to a substance in the sense of having genidentical parts, the ontological underpinning of spatiotemporal relations is more difficult to understand. The idea is that we are to think of spacetime points as enjoying a sort of fleeting existence—they are real, but they do not endure. These spacetime points are supposed to stand in spatiotemporal relations to one another. Furthermore, we should think of material events as *occurring at* given spacetime points; material events then stand in definite spatiotemporal relations to one another only in virtue of the relations between the spacetime points at which they occur.

3. So much for substantivalism about spacetime. How should we formulate four dimensional relationalism? This doctrine is even more elusive than its substantivalist counterpart. One problem is that it has seldom been explicitly advocated: Teller is the only author I know of who has offered an extended defense of this position (see pp. 380-83 of Teller 1991). Hints can be found elsewhere: Teller cites a very brief passage in §VI of Horwich 1978; §IV.1 of Friedman 1983 includes a discussion of relationalism which is neutral between the three and four dimensional versions; the formal work of Manders 1982 and Mundy 1983 is suggestive; DiSalle 1994 and Hoefer 1996 advocate positions which are quite similar to that of Teller, although both identify themselves as substantivalists. The position that I will develop here is closely related to those of Teller and others, although I believe that my presentation of it is more perspicuous.

I take it that the driving intuition behind relationalism about space is that one can maintain that bodies stand in spatial relations to one another—and even that there are determinate facts about the structure of possible but non-actual spatial relations—without being committed to the idea that these relations hold in virtue of relations amongst the members of an ontological substratum of existent points of space. Relationalists about spacetime should attempt to lift this intuition to the four dimensional context.

The most straightforward way of doing so is to say that the structure of spacetime is the structure of possible spatiotemporal relations between events, just as the structure of space is the structure of possible spatial relations between point particles. Relationalists will want to go on to insist that—just as in the three dimensional case—one can hold this view of spacetime without admitting into one's ontology a set of non-material points which stand in these relations, and which (via the notion of occupation) allow material events to stand in these relations. This is, of course, an obvious way to cash out the sort of relationalism in opposition to which Earman defines substantivalism (see proposal (iii) of the previous section).

Now, in the three dimensional case it seems that the relationalist can reasonably claim that the spatial relations between material bodies at a given time exhaust the physically meaningful information about the spatial situation of bodies at that time. A standard way of cashing this out is to remark that, in the three dimensional Euclidean case, specifying the spatial relations between particles fixes the imbedding of those particle into Euclidean space upto an isometry. It seems that the relationalist about spacetime is obliged to make the same sort of claim: that the spatiotemporal relations between events exhaust the physically meaningful information about the spatiotemporal situation of those events.

In some cases, such claims can indeed be made good. We can, for instance, regard Minkowski geometry as giving the structure of the possible spatiotemporal relations between bodies (within the domain of applicability of special relativity), since specifying the set of events along the worldlines of a set of particles and the Minkowski interval between each pair of events fixes the imbedding of the worldlines of the particles into Minkowski spacetime upto a Poincaré transformation (see §6.10 of Earman 1989 and §5 of Maudlin 1993). Thus one can regard the spatiotemporal relations between material events as exhausting spatiotemporal facts about bodies.[31]

Things do not work out so nicely in the non-relativistic case. Minkowski spacetime and Euclidean space are both semi-Riemannian geometries, so that all of their geometric structure can be encoded in a single object, the metric tensor, which gives the 'distances' between the points. Specifying the distances between things suffices to fix (upto isometry) the representation of the things as imbedded in the mathematical space/spacetime. Unfortunately, the same is not true of neo-Newtonian spacetime. Here one needs to specify a number of geometric objects—spatial metric, temporal metric, and affine connection—which do not, evenly jointly, give rise to a natural notion of distance between arbitrary pairs of points.

So it is not even clear what the four dimensional relationalist about neo-Newtonian spacetime should be hoping to prove. One cannot, of course, hope that specifying the spatial and temporal distances between events will fix their imbedding into neo-Newtonian spacetime, since this criterion ignores entirely the affine structure of the spacetime which is an essential part of its geometric structure. If we know, for instance, that two particles maintain a fixed distance from one another for all time, this does not suffice to fix an embedding into neo-Newtonian spacetime, since we do not yet know whether the particles are rotating or not. These metric data must, therefore, be supplemented by the inclusion of

[31] Throughout this Chapter I focus the discussion on bodies rather than fields. This is strictly for expository purposes, and makes no material difference to the conclusions that I reach.

some inherently affine ideology. Maudlin 1993 considers the possibility of including in the relationalist data "a three place predicate col(x,y,z) which has as its extension all triples of nonsimultaneous collinear events" (p. 193). So that col(x,y,z) iff it would be possible for a freely falling body to pass through each of x, y, and z. He then asks us to consider:

> two particles … that are uniformly rotating about their common center of mass. Until the first rotation is complete, no triple of occupied event locations are collinear. Even after any number of rotations, the collinearity relations among occupied points will be consistent with any periodic rotation, uniform or nonuniform. … Knowing, for example, that two particles remain at a constant distance through some period of time and that no three points on their worldlines are collinear, the neo-Newtonian relationalist could not predict when, if ever, a triple of points would be collinear. Nor could inertial effects be predicted or explained, since the absolute acceleration cannot be inferred from the data. (p. 194)

Maudlin takes this to be a definitive criticism of the program of four dimensional relationalism in the non-relativistic context.

I think that this conclusion, although natural from the perspective that we have been considering, is premature. At best, Maudlin's argument shows that the program of casting neo-Newtonian spacetime as the framework of possible spatiotemporal relations between events is in trouble if one tries to capture information about absolute acceleration using a relational predicate like col(x,y,z). But, of course, other options are available. One such option has already been discussed at length in Chapter 2: one can follow Sklar in regarding absolute acceleration as a primitive property of an event considered as an infinitesimal element of a worldline. This leads one to an interpretation of physics set in neo-Newtonian spacetime which is essentially the same as the three dimensional relationalist interpretation of the previous chapter.

Here I would like to sketch another interpretative possibility which is, I believe, a better candidate for a truly *four dimensional* relationalism.[32] The starting point is to return to the observation that neo-Newtonian spacetime is not a semi-Riemannian geometry. So one cannot regard it as representing the structure of possible distance relations between events. Thus far we have discussed the possibility of imposing additional non-metric

[32] I will limit myself to a mere sketch since a full elaboration of this line of thought would be a considerable undertaking. Here my main interest is with relativistic physics where the semi-Riemannian nature of geometry means that relationalists need not resort to such elaborate measures. Thus the following discussion is by way of an aside.

ideology on the ontology of events (such as col(x,y,z) or a primitive property of absolute acceleration). There is, however, another possibility: to shift the ontological focus away from events. Wordlines are the most promising alternatives—since their essentially four dimensional character suggests that they may be the appropriate ontology on which to base a relationalist account of spacetime.

I would like to suggest that relationalists should regard neo-Newtonian spacetime as representing the possible spatiotemporal relations between worldlines, rather than the possible spatiotemporal relations between events.[33] This is bound to strike many as a curious suggestion. Although there is nothing extraordinary about a geometry of lines (think of ordinary projective geometry), it is natural to ask what sort of spatiotemporal relations can hold between worldlines.

Part of the answer is obvious: if L and L' are worldlines, then we can speak of L and L' standing in certain relations in virtue of the relations in which the parts of L and L' stand. For example, if p and p' are events on L and L' respectively, then the fact that there is a certain temporal interval between p and p' (and a certain spatial interval as well, if p and p' are simultaneous) can be regarded as part of what it means to say that L and L' stand in a certain spatiotemporal relation to one another.

This does not get us very far, however, since it ignores the affine structure of neo-Newtonian spacetime. This affine structure underwrites a notion of the instantaneous acceleration of an event considered as being part of a worldline. It may seem that no respectable relationalist interpretation of neo-Newtonian spacetime can take advantage of this structure, since it doesn’t seem to make any sense to say that there could exist only a single worldline, and that that worldline could be everywhere accelerating. This is, of course, merely a restatement of the problem of absolute acceleration of Chapter 2. In the present context, this problem takes the following form. The spacetime substantivalist can make sense of the notion that if there were only a single worldline, it could be the worldline either

[33] This sort of program has recently been carried out in the case of GR; see Cassa 1996.

of a freely falling particle or of an accelerating particle—in one case, the worldline occupies a geodesic of physical spacetime, in the other it does not. But how can relationalists make sense of these two possibilities without an ontological underpinning which allows them to distinguish between inertial and accelerated motion?

I would like to sketch a relationalist reply to this objection, which will also serve as a partial explication of the notion that worldlines can stand in spatiotemporal relations to one another. In order to motivate this reply, it is helpful to return to an idea that was briefly discussed in Chapter 2: that relationalists about space can consistently maintain that there are as many kinds of empty worlds as there are possible geometries of space. Let's focus on a specific example. Imagine a non-flat Riemannian geometry on $\mathbf{R}^3$, which is such that reflection in the x-y plane is an isometry and that there is a scalar, f(x,y,z), constructed out of the curvature tensor which has its only local maxima at p=(0,0,1) and q=(0,0,-1). If we take this geometry to represent the unchanging geometry of physical space, then we can go on to add either a substantivalist or a relationalist gloss to our geometrical talk .

If we are substantivalists about space, then we will say that at any given time there are two points of physical space which have maximal curvature. These points maintain their identity over time, so that it makes sense to ask, for instance, if these points of maximal curvature at a given time are points of maximal curvature at other times. Similarly, it makes sense to describe two qualitatively identical possible worlds in which only one of the points of maximal curvature is occupied by a material object, such that the only difference between the worlds is which such point is occupied.

If, on the other hand, we wish to give a relationalist interpretation of this geometry, we will say that at a given time the abstract geometry represents the possible spatial relations between bodies. This network of possibilities is such that there are two distinguished possibilities, corresponding to p and q. They are distinguished in that they permit an object to stand in spatial relations to other objects which no other possibilities permit. In effect, the structure given to the set of possibilities by the scalar f gives the relationalist the resources to refer to individual 'parts' of space. This may make some suspicious that the distinction between substantivalism and relationalism is in danger of collapse. Indeed, we can imagine that the world under consideration is such that a set of curvature scalars gives us a coordinate system on space (upto reflection). So it seems that the relationalist is just as committed to the existence of the parts of space as the substantivalist is, since both recognize that these parts have characteristics which distinguish them from all the other parts of space.

I maintain, however, that even in this case relationalism is distinct from substantivalism. Relationalists will deny the substantivalist claim that it makes sense to imagine two qualitatively indistinguishable possible worlds which differ only as to which member of a pair of distinguished points is occupied. More importantly, they will not think that there is any sense to the question as to whether the geometrical properties of a given point change over time, since they deny that the parts of space maintain their identity over time. This, of course, is not an abstruse metaphysical point, since it means that relationalists have to work much harder than their substantivalist counterparts to introduce a notion of absolute motion. (On the other hand, substantivalists must work harder than relationalists in order to *avoid* introducing a notion of absolute velocity…)

So three dimensional relationalists can make sense of the idea that within the framework of possible spatial relations there are certain distinguished elements, without becoming crypto-substantivalists. This suggests a solution to the problem of acceleration for spacetime relationalists. The problem of acceleration looks like a formidable obstacle for spacetime relationalists only so long as we think of the geometry of possible relations between worldlines as being homogenous in the sense that each element of this geometry which corresponds to the existence of a single worldline is as good as as every other such element. For in this case, it will indeed appear to be impossible for relationalists to countenance the existence of distinct possible worlds, one of which contains a single freely falling particle, the other of which contains a single accelerated particle. But considering the three dimensional case should make us realize that relationalists can handle inhomogenous geometries. Thus, spacetime relationalists can handle the problem of absolute acceleration by stipulating that the space of possible spatiotemporal relations between worldlines is inhomogenous. In this way, they will be able to say "If there were only one worldline, it could be either inertial or non-inertial" just as three dimensional relationalists can say "If there were only one body, it could be at a location of maximum curvature, or not."

Furthermore, notice that four dimensional relationalists, just as much as their three dimensional counterparts, can distinguish themselves from substantivalists by looking at Leibniz-Clarke counterfactuals. Where the substantivalist sees many possible worlds containing a single uniformly accelerating particle (of a given mass, with a given acceleration), the relationalist sees only one possibility. This is the standard criterion for separating four dimensional relationalists from four dimensional substantivalists (see, e.g., Friedman 1983, p. 221 or Maudlin 1993, pp. 200-201).[34]

This apparently leaves us with two interpretative positions concerning neo-Newtonian spacetime, and with a criterion for distinguishing substantivalists from relationalists (i.e., the factions' respective stances on the Leibniz-Clarke counterfactuals). Not surprisingly, each of the positions under consideration is prima facie vulnerable to certain objections, and there is some question as to whether or not the criterion actually distinguishes between substantivalism and relationalism. In the next three sections I consider these objections in turn.

4. Against the substantivalism of §2, relationalists might be tempted to object that notion of spacetime points as existents is opaque. Existents, it might be maintained, are things which endure over time, and so it is impossible to understand what substantivalists mean to convey in declaring that they believe that spacetime points exist, beyond their stubborn refusal to admit that the considerations of Chapter 2 show that the substantivalist

[34] Note, however, that it no longer makes any sense to distinguish the doctrines via their stances on the persistence over time of the parts of space.

intuition is misguided. If substantivalists retort that spacetime points are to events as points of space are to objects, then it seems that they may indeed be in trouble since relationalists will likely retort that the notion of a pointlike event is intelligible only insofar as it is understood as being an infinitesimal segment of a worldline of an object which itself endures over time. Therefore, relationalists will claim, substantivalists are forced to admit that the notion of a spacetime point not so easy to explain after all. It would be a mistake to accede to this argument, however, since the notion of an existent which does not endure over time is to be found already within the ontology of relationalism about space in the form of the concept of a part of a field at a given instant. So there is no reason for substantivalists to be embarrassed about the idea that spacetime points enjoy a fleeting existence.

5. Substantivalists, on the other hand are liable (as usual) to charge relationalists with attempting to help themselves to all the advantages of the theoretical structure of neo-Newtonian spacetime without being willing to take it seriously when it comes time to interpret the theory. Substantivalists often speak as if the objections concerning absolute motion which were canvassed in the previous Chapter show that any sort of relationalism is doomed to failure. We have already seen, however, that relationalists about space can avoid this objection if they are willing to introduce a little extra ideology. In order to see how much pressure these considerations put on relationalism about spacetime, it is helpful to review the recent history of this line of thought. This history falls into two distinct phases: an early period, in which it was more or less taken for granted that the problem of absolute motion mitigated against any form of relationalism; and a more recent period, during which substantivalists have actively argued in favor of this claim.

As discussed in §2.7, Stein 1967 contains the seminal statement of the argument from absolute motion. This is often taken to provide a decisive argument in favor of substantivalism about spacetime. Stein himself, however, is not very interested in developing it in this way. He is, in fact, concerned to use his argument to deflate the substantival-relational debate rather than to resolve it. Thus he adopts a very ecumenical attitude towards the apparent conflict between Newton and Leibniz over the nature of space, attempting to emphasize their fundamental agreement concerning the status of space as an existent rather than their disagreement concerning the nature of this existence (pp. 193-94; see also Malament 1976, DiSalle 1994, and DiSalle 1995; see Rynasiewicz 1996a for critical discussion). Further, he seems to be skeptical that the modern substantival-relational debate has any content, once we have decided which spacetime to employ in formulating physical theories:

> If the distinction between inertial frames of reference and frames of reference which are not inertial is a distinction that has a real application to the world; that is if [Newtonian or neo-Newtonian spacetime] is in some sense [a structure] 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 the just indicated sense, this structure of spacetime also "really exists," surely seems to be supererogatory. (p. 193)

The emphasis here is on determining the structure of spacetime. Stein seems to be claiming that once this question is settled there are no additional interpretative questions of any substantive interest.

Earman 1970b marks a sharp departure from Stein's position. Earman, quite correctly in my opinion, emphasizes the disagreement between Newton and Leibniz while recognizing their points of agreement (pp. 288-89). Thus, at least in the case of space, he sees conceptual room for a disagreement over ontology even when there is agreement on structure. Furthermore, he sees Stein's argument concerning the relation between the problem of absolute motion and spacetime structure as being an argument in favor of spacetime substantivalism:

> for an account of motion, one needs to know not only the nature of the instantaneous spaces but also how these spaces, so to speak, fit together to give a spacetime structure, and the latter is surely not determined by the relations between material bodies (see [Stein 1967]). Moreover, we are inclined to hypostatize spacetime, to think of spacetime as something absolute, having meaning apart from the relations of bodies—how else is one to understand the notion of a worldline? (p. 313)

This paper, together with Earman and Friedman 1973, played a crucial role in establishing spacetime substantivalism as the dominant position in the philosophy of space and time.[35] I

[35] Note that even in Sklar 1974—where Sklar's Maneuver makes its first appearance—it is taken for granted that the burden of proof lies with the relationalists; see especially p. 165.

claim, however, that neither paper contains a cogent argument in favor of spacetime substantivalism over spacetime relationalism. Rather, they contain sophisticated developments of Stein's line of thought, explicating the role that inertial structure plays in non-relativistic physics. It is then taken for granted that this justifies an inference to the best explanation to a "realist interpretation of spacetime theories" (Earman and Friedman 1973 p. 329; see also p. 358)—and, furthermore, that this realism has more in common with the substantivalism of Newton and Clarke than with the relationalism of Leibniz.

Thus the early literature suffers from a blindness to the possibility of four dimensional relationalism. It follows that there was no appreciation of how there could be, as there was in the classical substantival-relational debate over space, agreement about structure but disagreement over ontology. This has been rectified, however, in light of the four dimensional re-interpretations of Sklar's Maneuver in Friedman 1983 and Teller 1991. Earman 1989 contains a concise statement of the standard argument given for preferring spacetime substantivalism over spacetime relationalism:

> The original Newtonian form held that [not all motion is relative], because the analysis of motion requires the absolute change of position, which concept has to be understood as change of position with respect to a substantival space. The modified form concedes that the scientific treatment of motion doesn't require absolute change of position or absolute velocity but asserts that it does require some absolute quantities of motion, such as absolute acceleration or rotation. To make these quantities meaningful requires the use of inertial structure or the like, and these structures must be properties of, or inhere in, something distinct from bodies. The only candidate for the role of supporting the nonrelational structure is the spacetime manifold… (p. 125)

In a similar vein, Hoefer 1996 lodges the following complaint against spacetime relationalism:

> I have difficulty understanding this view of the metric field [of GR] as not an existing thing, but a representor of free-floating dispositions; I prefer to see dispositions ascribed to concretely existing entities. For this reason, I think Teller's view is best described, in terms of the traditional debate, as a form of substantivalism. (p. 25)

As I understand Earman and Hoefer, they are claiming that the notion of a four dimensional spatiotemporal geometry is unintelligible unless it is recognized that the spatiotemporal structure is grounded in an ontology of existent spacetime points.

It should be clear by now that this is a very weak objection. Both substantivalists and relationalists make the transition from space to spacetime by straightforward generalization. Substantivalists move from a spatial geometry which is underwritten by existent points standing in geometric relations to a spatiotemporal geometry which is underwritten by existent points standing in geometric relations. Similarly, relationalists make the move from a spatial geometry conceived as the framework of possible spatial relations between bodies to a spatiotemporal geometry conceived as the framework of possible spatiotemporal relations between events or worldlines. And whereas the concern raised in the previous section constitutes a prima facie reason to be suspicious of the former transition, I can not see any grounds for objecting to the latter. If one admits that the relationalist conception of space is cogent, I do not see how one can balk at the relationalist conception of spacetime.

6. “If all the spatiotemporal relations between bodies were as they are, but every material thing were three metres to the left of where it is, it would be a different world.” Relationalists about either space or spacetime will want to claim that this Leibniz-Clarke counterfactual is false, or to deny that it has any truth-value at all. Prima facie, it seems clear that substantivalists will want to affirm the truth of such counterfactuals: if points of space or spacetime exist in a robust sense, then it seems that the fact that *this* point is occupied by *this* bit of stuff is one of the facts about our world, and it is perfectly cogent to suppose that it could have been otherwise while supposing all of the spatiotemporal relations between bodies to remain fixed, so long as corresponding changes are made in the occupation relations between all other points and bits of stuff.

Such counterfactuals provide a handy criterion for distinguishing substantivalists from relationalists—which seems especially valuable in the four dimensional case where the two opposing positions are quite obscure. To the best of my knowledge, no modern-day substantivalist showed any interest in denying the truth of the Leibniz-Clarke counterfactuals until the advent of the hole argument. As will become clear in the course of the detailed discussion of this argument in Chapter 7, denial of the general relativistic counterparts of the Leibniz-Clarke counterfactuals is mandatory if one is to avoid the conclusion that spacetime substantivalism about GR commits one to the conclusion that GR is an indeterministic theory. In the aftermath of the hole argument, a number of authors have claimed that substantivalists should follow Locke in denying *all* Leibniz-Clarke counterfactuals, including the non-relativistic ones.

Broadly speaking, such authors fall into to distinct categories. On the one hand, we have those who defend substantivalism by advocating purely *formal* responses to the hole argument (see especially Mundy 1992, Wilson 1993, and Leeds 1995). These authors maintain that one can decide whether or not a theory is deterministic by examining the mathematical structure of the theory, without looking at its interpretations. They claim, therefore, that the hole argument rests on some sort of mistake since interpretative issues are irrelevant to the question of determinism. Thus substantivalism cannot commit one to believing that GR is indeterministic. The status of the Leibniz-Clarke counterfactuals within these formal approaches is somewhat difficult to determine. On the one hand, it would seem that these authors should be agnostic, since they are explicitly attempting to avoid engagement with interpretative issues. But Mundy, at least, seems committed to the view that all such counterfactuals are either false or senseless. This apparently stems from his deep skepticism concerning the relevance of modal language to philosophy of physics. In any case, these formal approaches will be shown to be untenable in Chapter 6, so we can set them aside in considering objections to using the non-relativistic Leibniz-Clarke counterfactuals as a criterion for distinguishing between substantivalists and relationalists.

There is, however, a second group of authors who espouse substantivalism but deny the Leibniz-Clarke counterfactuals (see Butterfield 1987, Butterfield 1989b, and Butterfield 1989a; and Maidens 1993). Both of these authors blame the fact that substantivalists have traditionally accepted the Leibniz-Clarke counterfactuals on the uncritical acceptance of a simplistic account of modal semantics.

Their argument runs as follows. Consider how the connection between the Leibniz-Clarke counterfactuals and substantivalism is usually motivated. One considers a possible world W whose ontology is made up of spacetime points and a single particle. We suppose for definiteness that the spacetime of W has the structure of neo-Newtonian spacetime and that the particle is freely falling. Now, there is a single one dimensional family, L, of spacetime points in W which are occupied by matter. Next we imagine a putative possible world, W', with the same structure, spacetime points, and particle as W, but in which the family of occupied points is some timelike geodesic $L' \neq L$. Now it seems that if W is a possible world, then so is W'. Furthermore, it is clear that the substantivalist must recognize that W and W' are *distinct* possible worlds, since if $p \in L$ but $p \notin L'$, then p is an existent which has different properties in the two worlds (in W but not W', p is occupied by a material event). It follows that substantivalists must endorse the Leibniz-Clarke counterfactual "if the particle had been *here* rather than *there*, the world would have been different."

Butterfield and Maidens refuse to accept the premise that there is a $W' \neq W$ with the same spacetime points as W. That is, they deny that there is a primitive notion of identity of spacetime points across possible worlds. This denial undercuts the argument, since it depends crucially on the claim that there is a point which has different properties in the two worlds. It is necessary to examine separately the grounds that these authors offer for rejecting the trans-world identity of spacetime points.

Butterfield follows Lewis in claiming that there is *no* trans-world identity (see Lewis 1973 and Lewis 1986). If it is true that this dissertation could have been better if only I had worked harder, it is not because there is a possible world where *I* worked harder and wrote a better dissertation. Rather it is because there is a possible world where someone very much like me, my *counterpart*, worked harder and wrote a better dissertation. According to Lewis, it is qualitative similarity which settles whether or not a denizen of another possible world counts as my counterpart in a given situation. The degree of similarity which is demanded varies from context to context: in an ordinary conversational context it will be true that if I had worked harder, I would have written better; but in a context where my personality and motivation are considered to be fairly rigidly fixed, it may not be true that we can find a possible world with someone sufficiently similar to me who worked harder. Butterfield's idea is that if we are given two possible worlds, each of which is composed of a spacetime with neo-Newtonian structure and a single freely falling particle, the most natural counterpart relation to choose for construing the Leibniz-Clarke counterfactuals is the one which agrees on *all* qualitative properties—i.e., the counterpart relation which makes the occupied points of one world the counterparts of the occupied points of the other. Thus the Leibniz-Clarke counterfactuals are false. This is an attractive consequence since it apparently blocks the hole argument—although it leaves one in disagreement with the classical substantivalists.[36] But, as Brighouse 1994 points out, this natural counterpart relation is merely one of many possible counterpart relations. It is true that it is privileged in that it is the one most faithful to the qualitative properties of things, but Lewis explicitly allows that there are contexts in which counterpart relations may match up things which share relatively few qualitative properties (see especially Lewis' discussion of situations in which *any* person could be considered to be his counterpart, on pp. 231-32 of Lewis 1986). In particular, there is nothing in Lewis' account to prevent one from counting an unoccupied point as the counterpart of an occupied point. Such a counterpart relation would, of course, make the Leibniz-Clarke counterfactuals true. Thus the adoption of Lewis' counterpart-theoretic modal semantics does not by itself prevent substantivalists from accepting the Leibniz-Clarke counterfactuals.

Maidens is willing to accept some notion of trans-world identity. But she suggests that our ideas about trans-world identity must be informed by our best physical theories. As an example, she discusses the Cantabridgian position that the strange statistics of many

[36] The success of Butterfield's response to the hole argument will be evaluated in the Appendix. For now: notice that there is reason to be skeptical, since Butterfield's counterpart-theoretic response works only if the natural counterpart relation (and no other) is used in determining whether or not a theory is deterministic.

particle quantum mechanics show that quantum particles cannot be identified across possible worlds in the same way that dice, say, can be (see Redhead 1990, Butterfield 1993, and Teller 1995). In particular, while it seems to be quite sensible to say of a pair of dice that although this one came up six and that one came up one, the results could have been reversed, one gets into trouble if one says of a pair of electrons that although this one has spin up and that one has spin down, it could have been the other way around—since the acceptance of latter sort of counterfactual, together with some form of the Principle of Indifference, entails that electrons should obey Maxwell-Boltzmann statistics rather than Fermi-Dirac statistics. Something has to go. A standard solution to this problem is to maintain that dice, but not electrons, possess a sort of "transcendental label identity" which supports trans-world identification.[37] Here we have an example of physics informing metaphysics. Maidens' suggestion is that the true lesson of the hole argument may be that one has to be just as careful with trans-world identification of spacetime points. If one naively posits two qualitatively indistinguishable possible worlds, in one of which point p is occupied and q is unoccupied, while in the other the situations are reversed, one runs afoul of the hole argument. The moral is not that spacetime points don't exist but that spacetime points, like electrons, lack transcendental label identity.[38] Since Maidens is willing to carry this conclusion over into the non-relativistic case, she will deny the Leibniz-Clarke counterfactuals. This strikes me as being an ingenious and successful response to the hole argument. But I think that it is important to emphasize that it is a desperate one—one must be deeply committed to the substantivalism in order to think it worthwhile to assimilate the semantics of spacetime points to the deeply mysterious and highly contentious semantics of quanta.

There are undoubtedly other grounds upon which substantivalists could follow Locke in rejecting the Leibniz-Clarke counterfactuals.[39] But I think that the shortcomings of

[37] Of course, we could have chosen to jettison the Principle of Indifference instead.
[38] In the case of the hole argument it is difficult to find other premises to deny.
[39] Hoefer 1996, like Butterfield, categorically rejects the notion of transworld identity, although, unlike Butterfield, he does not provide a principled account of our modal

the proposals of Butterfield and Maidens suggest an interesting dilemma for any substantivalist who wishes to do so.

On the one hand, one can, like Butterfield, offer a general account of modal semantics which has as a consequence the proposition that Leibniz-Clarke counterfactuals are false. But then one is faced with the unpleasant fact that in our ordinary discourse, we accept as true counterfactuals about dice, etc., which are the counterparts of the Leibniz-Clarke counterfactuals. Thus one's general account of modal semantics must have the resources for distinguishing between different sorts of Leibniz-Clarke-style counterfactuals, based on the nature of the entities under consideration. I do not think that one can in general hope to find such resources—certainly Lewis' counterpart theory does not contain them. On the other hand, one can follow Maidens in focusing from the beginning on the special nature of spacetime points, and arguing that they deserve exceptional treatment in our modal semantics. This is a viable strategy, but it means departing from the original substantivalist intuition that spacetime points are ordinary existents. Notice that in either case, one is driven to do a lot of hard work in order to be a substantivalist. One can't help wondering whether substantivalism is worth the effort.

Lockean substantivalists must also face the criticism that their rejection of the Leibniz-Clarke counterfactuals threatens to dissolve the substantival-relational opposition. The only apparent difference between a spacetime relationalists and a Lockean substantivalist about spacetime is that the latter insists that spacetime points exist, although this existence does not seem to have many consequences. This criticism will be made much more precise in Chapters 7 and 8.

I conclude that it *is* possible for substantivalists to reject the Leibniz-Clarke counterfactuals. But as in Locke's day, it requires quite a bit of effort. So in what follows, I

---

discourse. Hoefer's idea is that in the absence of primitive transworld identity, it does not in general make any sense to say of two things A and B that A might have been just like B and B just like A. This will indeed block the Leibniz-Clarke counterfactuals. I will not here attempt to discuss the details of Hoefer's theory of identity since it is, by his own admission, incomplete (see p. 22).

will continue to use the Leibniz-Clarke counterfactuals as a crude criterion for separating substantivalists from relationalists. However, when it is important to be scrupulous—as in c 7—I will take into consideration varieties of substantivalist who deny the Leibniz-Clarke counterfactuals.

7. This brings us back to the position we were in at the end of §3: we have on the table versions of four dimensional substantivalism and relationalism and a rough criterion for distinguishing the two factions. I, for one, am confident that both interpretations are coherent. Furthermore, I claim that we have no decisive grounds for preferring one or the other.

In Part II of the dissertation, I will be considering ways in which interpretative questions about quantum and classical physics can interact, and will argue that taking the conceptual and technical problems of quantum gravity into account can cause us to rethink our interpretative judgments about GR. For the time being, however, I propose to set aside such considerations. And it seems clear that if we limit ourselves to classical physics, we can find no physical ground for preferring one interpretation over the other. Since the four dimensional substantivalism and relationalism that we are considering are both interpretations of the full structure of neo-Newtonian spacetime, we cannot hope to distinguish between them on any physical grounds arising out of any theory which can be set in neo-Newtonian spacetime. In particular, both interpretations support exactly the same notions of absolute motion, so we cannot criticize one or the other for allowing too many or too few such notions.

There are, of course, metaphysical grounds for preferring one or the other theory. It is generally granted that it is an advantage to be able to deny the Leibniz-Clarke counterfactuals (see, e.g., Friedman 1983, p. 221). Although it is universally recognized that Leibniz's arguments against the existence of absolute space are not conclusive objections unless one grants the truth of the Principle of the Identity of Indiscernibles or the Principle of Sufficient Reason, nonetheless the radical multiplicity of qualitatively indistinguishable possible worlds which is implied by the truth of the Leibniz-Clarke counterfactuals is held to be undesirable in itself. Thus relationalism has a distinct advantage over substantivalism, to the extent that commitment to the latter implies commitment to the truth of the counterfactuals.

On the other hand, I suspect that substantivalists will continue to insist that their account of the nature of spacetime has explanatory virtues which the relationalist account lacks. They may, for instance, concede that it is possible to conceive of affine structure as being ontologically free-standing, but suggest that it is more satisfying to conceive of it as being the structure *of* a set of spacetime points. Similarly, they may continue to call attention to the putative explanatory strengths mentioned in the previous chapter (see §§2.4 and 2.9).

In the end, the issue comes down to questions of taste of the most abstruse sort. Which interpretation has the more intelligible ontology? Which provides the most satisfying account of the modality of spatiotemporal properties? At this point I am again inclined to declare an impasse. I do not have strong enough feelings about these questions to prefer one interpretation over the other. So I am satisfied to allow that both interpretations are viable and that they are, modulo some very fine points of metaphysics, more or less equally attractive.

8. Finally, we are left with the question of the relative merits of three and four dimensional interpretations. Here I believe that the situation is slightly more clear-cut.

One the one hand, it seems to me that the three dimensional interpretations have a clear advantage in terms of intelligibility. This is especially true in the case of substantivalism, where the doctrine about space is quite intuitive and the doctrine about spacetime is quite difficult to grasp—indeed whatever intelligibility the latter has seems to be parasitic on that of the former. In the case of relationalism, the disparity is not so great, particularly if we are discussing a four dimensional relationalism which construes spacetime as providing the structure of possible relations between events rather than between worldlines. But even here, I think that most people will agree that it is much easier to grasp the idea of a geometry of relations between objects than a geometry of relations between events.

On the other hand, it seems clear that the four dimensional interpretations provide much more satisfying accounts of absolute motion. Again, the advantage is greater in the case of substantivalism, since the substantivalist about spacetime entirely avoids the embarrassment associated with the three dimensional substantivalist's acceptance of absolute velocity. Even in the relationalist case, it seems that the transition to four dimensions brings a significant advantage: where the relationalist about space was forced to accept a somewhat mysterious looking primitive property of absolute acceleration, the spacetime relationalist is able to assimilate the problem of absolute acceleration to the general problem of providing a relational account of geometry.

I am inclined to see a decisive advantage for the four dimensional approach here. Particularly, of course, for spacetime substantivalism where the gain in the physically grounded problem of absolute motion has to more than outweigh any (strictly metaphysical) disadvantage of relative intelligibility. But even in the case of relationalism about space, I think that the gain in elegance over the three dimensional doctrine is well worth the loss in intelligibility. (I am also, of course, willing to concede that in the case of the latter judgment I have strayed into questions of metaphysics…)

CHAPTER 4
General Relativity and the Substantival-Relational Debate

1. The project of this first Part of the dissertation is to evaluate the arguments which have been given in favor of the received view that relationalism concerning the spacetime of GR is untenable. We are now in a position to complete this project. This turns out to be a relatively straightforward task, since most of the hard work has already been done in Chapters 2 and 3. In the next section, I will briefly discuss the prospects of the three dimensional interpretations in the general relativistic context. In the following sections, I turn to my main business in this chapter, which is to argue that four dimensional relationalism is a viable alternative to spacetime substantivalism.

2. I do not know of any philosopher who entertains, let alone advocates, substantivalism about space as an interpretative option for GR. Relationalism about space, on the other hand, is sometimes briefly considered—only to be rejected. Typically what happens is that a relationalist is posited who is saddled with an ontology of point particles, and is assigned the patently impossible task of predicting the future behavior of the particles from knowledge of their positions and velocities at a given time. One then observes that, in addition to the obstacles which the program of Machianization faces in the Newtonian context (see §§2.10-12), this sort of relationalism about GR is made impossible by the fact that the relationalist data may not determine the geometry of space at a given time, so that there will in general be many different physically possible futures for a given set of data (e.g., gravitational waves may or may not affect the particles in the future; see §6.10 of Earman 1989 and §6 of Maudlin 1993). This sort of relationalist is, of course, a straw man who is set up only to make a point about the dynamical nature of the spacetime of GR. To the extent that relationalism about GR is discussed, it is a variety of spacetime relationalism which is considered. Thus neither sort of three dimensional interpretation is taken seriously.

This is likely to seem quite reasonable—after all, wasn't Minkowski warranted in proclaiming that "Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality" (Minkowski 1952, p. 75)? The manifest moral of relativistic physics is that there is no preferred notion of simultaneity, so that we must take spacetime as our fundamental notion and regard space and time as (at best) derivative—"talk about space and time is really talk about the spatial and temporal aspects of space-time" (Earman 1970a, p. 259).

So long as we read the physical content of GR along these lines, three dimensional interpretations will not be very attractive, since they require that we recognize within each model of the theory a preferred slicing into hypersurfaces of simultaneity.[40] Furthermore, of course, some models (such as Gödel spacetime) do not admit a foliation by spacelike surfaces. Even among models which do admit such foliations, some models will not admit a preferred slicing, since their geometry does not provide the resources necessary to single out a particular slicing. In Minkowski spacetime, for instance, one cannot distinguish between a given slicing and its Lorentz boost using only predicates which are definable in terms of geometric structure. Of course, some models *do* permit a preferred slicing. Let's say that a model admits a CMC slicing if it can be foliated by surfaces of constant mean curvature (the mean curvature is just the trace of the extrinsic curvature; see §8.4). Within the space of solutions of the Einstein field equations there is an open set (of unknown extent) of models which admit a CMC slicing (see Isenberg and Moncrief 1996 and the references therein). Many of these models admit a unique such slicing. One does not, however, want to jump to the conclusion that GR does, after all, countenance a preferred notion of simultaneity. For one thing, models which admit a preferred CMC slicing may also provide the resources for constructing other distinguished slicings. More importantly, GR itself does not suggest any *physical* reason to regard the existence of a unique CMC slicing as revealing any deep fact about the nature of a given spacetime.

[40] Bohmians suggest that accepting a three dimensional interpretation is the key to making sense of quantum theories; see, e.g., Cushing 1996 and Weingard 1996. Most commentators regard this as a serious shortcoming of the Bohmian approach; see, e.g., Maudlin 1994, pp. 212-17 and 234.

Thus GR does not seem to provide any grounds for viewing space and time as fundamental realities.[41] As a consequence, the three dimensional interpretations of GR have not seemed very interesting, and have not been discussed in the philosophical literature on the substantival-relational debate. In Chapter 8 I will argue that the difficulties associated with the notions of time and observables in quantum gravity can motivate us to rethink this conclusion. In short: one way to solve these problems is to stipulate that the CMC slicing has genuine physical content; accepting this solution leads to a view of GR as a theory of the evolution in time (as measured by mean curvature) of the geometry of space (relationally conceived). Thus considerations arising out of quantum gravity may reveal a hidden moral implicit in GR, which undercuts the significance of the manifest moral.

3. I will take it for granted that the discussion of Chapter 3 suffices to show that spacetime relationalism and spacetime substantivalism are viable interpretative strategies in the Newtonian context. My goal here is to argue that, upto the problems to be considered in Part II, the transition to GR does not alter this situation. That is, if we attend only to the concerns addressed in the substantival-relational literature prior to Earman and Norton 1987 then we find no decisive grounds for preferring one of the other sort of four dimensional interpretation.[42]

There are three ways in which GR differs from Newtonian gravity set in neo-Newtonian spacetime which seem as though they might be pertinent to the substantival-relational debate.

(1) In GR the geometric structure is encoded in a metric of Lorentz signature rather than in the familiar spatial, temporal, and affine structure of neo-Newtonian spacetime. This means

[41]Barbour 1994a comes closer than anyone to denying this claim (see especially §§1 and 2). Ultimately, however, he seems to rely upon developments in quantum gravity to justify his view.

[42] That is, we bracket intertheoretic concerns, and intratheoretic concerns related to the general covariance of GR. We focus on problems related to the existence of absolute motion and the field-theoretic nature of contemporary physics

that there is no absolute notion of simultaneity, and that the geometric relations between two events are exhausted by the facts about the lengths of curves connecting those events.

(2) This geometric structure is dynamic. That is, the geometric structure varies from model to model. This variation is in part due to the dependence of the metric on matter distribution. But it is important to emphasize that, according to GR, there are many possible spatiotemporal geometries even when no matter-energy is present.

(3) GR is first and foremost a field theory. In the Newtonian context one can represent matter using either point particles or fields. Indeed, one can do quite a bit of physics without mentioning fields at all. If, for instance, one is interested in studying systems of gravitating bodies, one can often treat problems in full detail while representing the bodies as point particles acting upon one another at a distance. In GR, one still uses both the field-like stress energy tensor and point-particle test bodies to represent matter. But there is a sense in which fields are the primary representors of matter. One can, of course, perform quite useful calculations—such as finding the perihelion of Mercury to within observational error—by examining the behavior of test particles in spacetimes with vanishing stress-energy tensors. But if one wanted to treat a gravitational problem involving even two bodies in full detail, one would have to represent those bodies via a non-vanishing stress-energy tensor. Thus, in GR the postulation of fields is no longer optional.

It is convenient to consider the complications introduced by these three features of GR in sequence rather than simultaneously. Therefore, in the next section I will consider the interpretative possibilities for Special Relativity (SR), where only (1) is in effect; in §5 I will consider the case of vacuum GR where (1) and (2) must both be considered; in the final section, I will turn to the case of full GR, where (1), (2), and (3) must all be taken into account.

4. Special relativistic physical theories are set in Minkowski spacetime. Our task here is to confirm that there are viable four dimensional substantivalist and relationalist interpretations of such theories.

The transition to Minkowski spacetime is straightforward for substantivalists. They can simply maintain that physical spacetime is composed of existent points which stand in determinate spatiotemporal relations to one another; that the structure of this spacetime is Minkowskian; and, furthermore, that material events stand in spatiotemporal relations to one another in virtue of the relations which hold between the points which they occupy.

For relationalists, on the other hand, the transition from neo-Newtonian spacetime to Minkowski spacetime leads to a slight shift of position. Recall from §3.3 that the degenerate nature of the spatial and temporal metric structure of neo-Newtonian spacetime made it impossible for relationalists to account for the existence of absolute acceleration while construing spacetime as the framework of possible spatiotemporal relations between events. Thus the problem of absolute motion drove relationalists to regard spacetime as being the (inhomogeneous) geometry of possible spatiotemporal relations between worldlines. This complication is unnecessary in the domain or SR, where the semi-Riemannian nature of the metric of Minkowski geometry means that metric relations exhaust geometric content, so that specification of the spacetime intervals between pairs of events determines their imbedding into Minkowski spacetime upto a Poincaré transformation. For this reason, it is generally allowed that four dimensional relationalists about SR can claim that spacetime is just the set of possible spatiotemporal relations between events, just as three dimensional relationalists about Newtonian spacetime can claim that Euclidean space is just the set of possible spatial relations between bodies.

Thus we have two putative interpretations of Minkowski spacetime. I claim that both of these interpretations are viable, upto problems of metaphysics. Indeed, for the time being we are setting aside the problem of fields, and (as in Chapter 3) the problem of absolute motion cannot help us to distinguish between the two interpretations (since they recognize exactly the same geometric structure, and hence underwrite exactly the same notions of absolute motion). Furthermore, just as in Chapters 2 and 3, relationalists and substantivalists will disagree about the status of the Leibniz-Clarke counterfactuals.

5. We now move on to consider vacuum GR, where we must contend with a Lorentzian metric which is dynamic. Here substantivalists will claim that a given solution, (M,g), of the vacuum Einstein field equations (EFE) represents a physical spacetime whose existent points stand in spatiotemporal relations isomorphic to those which hold between the points of M. The only complication which the move to GR introduces is that substantivalists must now be very careful about Leibniz-Clarke counterfactuals. In the Newtonian context, it seemed natural for substantivalists to say that two possible worlds could be qualitatively indistinguishable but distinct, as is true of the pairs of worlds involved in Leibniz-Clarke counterfactuals. We will see in Chapter 7 that adoption of this attitude towards the worlds represented by models of GR leads to the conclusion that the theory is indeterministic. On the other hand, as discussed in §3.6, rejection of the counterfactuals is not an entirely straightforward task.

Of course, this is not a problem for relationalists, who were never tempted to affirm the Leibniz-Clarke counterfactuals in the first place. But they are faced with their own difficulties in moving from the special case of Minkowski spacetime to the full class of vacuum solutions to the EFE. In former case, it was unproblematic to maintain that spacetime was the set of possible spatiotemporal relations between events, because we knew what we meant by 'spatiotemporal relations'—the spacetime interval between pairs of events. But, as Earman points out, the situation is less straightforward in the case at hand, since in a generic general relativistic spacetime "an arbitrary pair of events may not be connected by a geodesic, and even if it is, the geodesic may not be unique" (Earman 1989, p. 130). Thus it seems that relationalists will have difficulty in specifying the set of spatiotemporal relations of which spacetime is supposed to be the framework of possibility.

This objection is inconclusive, however. Recall that every model (M,g) of GR is locally convex , in the sense that every point p∈M has a neighborhood, U, such that every pair of points in U is connected by a unique geodesic (see Theorem 8.1.2 of Wald 1984). This means that, locally, relationalists can treat a vacuum solution as they would Minkowski spacetime, by asserting that the spacetime is just the set of possible geodesic distances. Globally, relationalists can regard spacetime as being a patchwork of overlapping convex neighborhoods, so that spacetime is still regarded as being the framework of possible spatiotemporal relations between events, although 'spatiotemporal relation' now means 'piecewise geodesic distance through a particular series of convex neighborhoods.'[43]

This establishes that we can generalize the argument of §4, and conclude that the Lorentzian nature of the metrics of the spacetimes GR does not by itself present any obstacle to relationalism. We now have to examine the interpretative consequences of the fact that spatiotemporal geometry is dynamic in GR.

As was mentioned in §3.2, it is occasionally maintained that the incompatibility of relationalism and vacuum GR follows immediately from the fact that the latter countenances the existence of a multiplicity of distinct matter-free possible worlds. This view is motivated, presumably, by the intuition that only those who liken spacetime to a substance can grant it this robust sort of subsistence. In a similar vein, Maudlin argues that in order to avoid the objection to a relationalism founded on a point particle ontology which was mentioned in §2, relationalists about GR must subscribe to a plenist ontology. From the fact that no physical field is non-zero in vacuum GR, it supposedly follows that relationalists must recognize the metric field itself as an existent. But "[i]f relationists are willing to admit into their ontology a metric field, which fundamentally has only spatiotemporal properties, as a substance, why not call it spacetime and join the ranks of substantivalists?" (Maudlin 1993, p. 202).

[43] Note that relationalists about space may have to resort to the same ploy if they are dealing with a spatial geometry, such as the surface of a sphere, in which pairs of points are not always connected by a unique geodesic.

Against the first sort of objection, relationalists can reply as follows. As discussed at length in §2.4, relationalists about space maintain that space exists, and are free to go on to insist that it subsists, in the sense that it would exists even in the absence of matter. Having done so, there is no reason—other than commitment to a quite strong, and thoroughly optional, supervenience thesis—why they cannot go on to endorse the possibility of various sorts of matter-free space, distinguished from one another in virtue of providing distinct geometries of possible relations between bodies. The same series of moves is available for relationalists about the spacetime of GR. Thus, relationalists of all persuasions are free to endorse the subsistence of space/spacetime.

As for Maudlin's complaint, relationalists should deny that they are committed to regarding the metric field as a substance. All that they are committed to is regarding spacetime as an existent which can have a number of possible structures (which may or may not be distinguishable by local measurements at a given time, and which may lead to quite different long-time behavior of test-particles). But, of course, taking spacetime to be an existent is very different from taking it to be a substance. And recognizing that there is a multiplicity of possible structures for empty spacetime in no way commits one to the latter. Furthermore, all of this applies, *mutatis mutandis*, to relationalism about space—although this has not always been easy to see due to the traditional preoccupation with Euclidean space (on this point, see Nerlich 1991).

Similarly, it is sometimes suggested that the existence of gravitational waves in GR is in some way incompatible with relationalism. I believe that the intuition here is that if a gravitational wave flattened me, then I would admit that spacetime itself had causal powers. Somehow, this is supposed to be tantamount to substantivalism. But, of course, if this counts as having causal powers, then Newtonian spacetime also has causal powers—since it too has the potential to flatten me if I struggle against its affine structure. Thus, again, the move to a dynamical spacetime does not affect the interpretative status quo.

In the context of vacuum GR, then, the dynamic nature of spacetime amounts to the fact that there is a multiplicity of matter-free solutions to the EFE. And relationalists are free to embrace each of these solutions as representing a genuine physical possibility for a way the world could have been had there not been any matter.[44] Thus they, like their substantivalists counterparts, will read the EFE as providing constraints on the structure of spacetime in the absence of source fields. Since we again have a situation where substantivalists and relationalists recognize equal amounts of spatiotemporal structure, the problem of absolute motion cannot give us any grounds for preferring one or the other interpretative strategy. So we once again have a situation where, modulo certain metaphysical considerations, substantivalism and relationalism appear to be equally attractive. The only novelty is that these considerations now include differences of opinion concerning the status and significance of the Leibniz-Clarke counterfactuals.

6. Finally, we turn to full GR, in which the stress-energy tensor is allowed to be non-zero. Here we are interested only in solutions which are physically realistic in the sense that the stress energy tensor is the stress energy tensor *of* physical fields (such as the electromagnetic field, the Klein-Gordon field, or phenomenological fields representing dusts or fluids, etc.). For simplicity, I will assume that we are talking about a single such field.

Let's begin by treating the field as a test field—that is, we will assume that it is defined on some background spacetime, but that it does not contribute to the stress energy tensor. Substantivalists and relationalists will both look to their three dimensional treatments of fields (see §§2.5 and 2.6). Substantivalists will maintain that a field is just an assignment of properties to existent spacetime points. Relationalists about spacetime will treat fields as being spatiotemporally extended objects whose infinitesimal parts are events which stand in

[44] Relationalists should, as in Chapters 2 and 3, make sense of this sort of talk in terms of counterfactuals—what would the world have been like if there had been a single particle? Here there may seem to be a difficulty, since adding even a single particle would change the spatiotemporal geometry. But this can be finessed—either by appealing to the familiar fiction of test particles, or by gesturing towards an account of the sense in which the vacuum geometry approximates various one particle geometries.

spatiotemporal relations to one another, just as relationalists about space view fields as being spatially extended objects with parts which stand in determinate spatial relations to one another. Thus both parties are able to handle test fields without difficulty.

Now it is a small step to allow the field to influence the geometry of spacetime, via its stress-energy and the EFE. For both parties, the constraints that GR places on physical possibility are a matter of delicate accommodation between matter-energy and geometry. For the substantivalist, points have both geometric properties and field properties, and these must be carefully assigned so that both the EFE and the equations governing the physical field hold. For the relationalist, the parts of a field stand in geometric relations to one another, as well as having certain properties which satisfy the field equations; but these geometric properties are intricately inter-related with the field strengths. Under both readings, GR is a marvelous theory. But there is no conceptual obstacle standing in the way of either interpretative strategy, so far as I can see.

We once again have a situation where substantivalists and relationalists recognize exactly the same amount of structure, so that the two camps are equally immune or vulnerable to the problem of absolute motion. And both are, of course, immune to the problem of fields. We are left, then, with the familiar conclusion that spacetime relationalism and substantivalism are equally attractive, upto certain considerations. As in the preceding discussions, these considerations include abstruse points of metaphysics. In the present case, however, we must also include differences of opinion over the status of the Leibniz-Clarke counterfactuals, and related issues concerning the identification of spacetime points across possible worlds. In Part II it will become clear that these considerations are closely related to the interpretative problems of quantum gravity. So I will end this part by concluding, not that four dimensional substantivalism and relationalism are both viable interpretations of GR, but merely that both should be considered to be live options as we head into Part II.

PART II
Gravity and Gauge

CHAPTER 5
Gauge Theories: Formalism

1. Introduction.

In Part I, I surveyed the arguments of the traditional substantival-relational debate, and concluded that neither faction has offered decisive grounds for preferring their approach to interpreting GR. Although both camps claim some advantages over the other—deriving from a supposed superiority of intelligibility, economy, or explanatory power—neither side is able to offer more than *metaphysical* grounds for preferring their interpretative approach. In this second Part, I will argue that the interpretative interplay between classical and quantum gravity may reveal *physical* grounds for choosing an interpretation of GR: since substantivalism and relationalism are associated with (mandate/suggest/are suggested by/are mandated by) distinct approaches to the outstanding conceptual problems of quantum gravity, commitment to a theory of quantum gravity may imply (or constrain) a choice between substantivalism and relationalism.

It will not, in the end, be difficult to establish that there is indeed such interpretative interplay between classical and quantum gravity. But in order to approach this problem we first need to develop a framework which is appropriate both for discussing the interpretative problems of GR and for approaching the problems of quantum gravity. I will limit my discussion of quantum gravity to a single approach: canonical quantum gravity (see Isham 1994 for a survey of alternative approaches). In the canonical approach to quantization, one begins with a Hamiltonian formulation of a classical theory and—with luck—constructs a quantization by finding an appropriate Hilbert space representation of the algebra of observables corresponding to the canonical variables of the classical phase space. Thus my first task in this second Part of the dissertation is to explicate the interpretative problems of GR within the Hamiltonian framework. From this point of view, GR is what is known as a *gauge theory*—the initial value problem for the EFE (Einstein field equations) is not well posed, since there are many solutions for each set of initial data, so that GR is usually thought of as having more variables than physical degrees of freedom.[45] This introduces some complications into the formulation and interpretation of GR in Hamiltonian form. I choose to approach these complications by first discussing the structure and content of gauge theories from a general point of view, before dealing with the additional problems specific to GR. This Chapter and the next are given over to an elaboration of the formalism and interpretation of gauge theories, respectively. In Chapter 7, I present GR as a gauge theory, and discuss its interpretations. It is at this point that the hole argument of Earman and Norton 1987 comes to the fore. I conclude that it is only with great difficulty that one can avoid the hole argument's conclusion—that GR is indeterministic when given a substantivalist interpretation—but that since this indeterminism has no empirical consequences at the classical level, this is not by itself a sufficient reason to reject substantivalism. Finally in Chapter 8, as promised, I provide an introduction to the central conceptual problems facing canonical quantization of gravity, and explore the ways in which some proposed solutions are associated with interpretative stances towards the classical theory. This leaves the outcome of the substantival-relational debate GR dependent on the

[45] 'Gauge' is used in a number of ways in the literature. Here I have opted for a sense which is quite precise and of intermediate generality. For my purposes, a gauge theory will be a constrained Hamiltonian system—this is very similar to an ordinary Hamiltonian system, except that the equations of motion contain arbitrary functions of time. There will be a number of terms—'gauge freedom,' 'gauge orbit,' 'gauge transformation'—whose sense will derive from this use of 'gauge theory.'

There are, however, several other perfectly consistent conventions for the use of the term. Under the most narrow convention, only Yang-Mills style theories of connections on principal fiber bundles count as gauge theories; in this case, GR, although a constrained Hamiltonian system, is not properly speaking a gauge theory (see, e.g., Duval and Künzle 1984 for this usage). At the other extreme, any freedom in describing a physical system is called a gauge freedom. Marsden, for instance, characterizes the freedom available in selecting a Lagrangian for *any* given classical system as a reflection of gauge invariance (see p. 31 of Marsden 1992). Some caution is required, since authors often move back and forth between the various senses of 'gauge'; e.g., Kuchar 1980 employs the term in three distinct senses which the reader must be able to distinguish.

details of quantum gravity and its interpretation rather than on abstruse points of metaphysics.

My goal in this Chapter is to present an exposition of the formalism of gauge theories. I attempt to steer clear of all issues of interpretation, although I often use the standard interpretation of the phase space of classical mechanics as representing the space of possible positions and momenta of a physical system to motivate the discussion. In presenting this material I have opted for a more or less coordinate free approach, since this allows for the most compact presentation. It has the advantage over more coordinate-based approaches of presenting the material in a very clean and geometric fashion, so that the roles of the various structures in play are very clear. Of course, it also has the disadvantage of being somewhat abstract, and far removed from the familiar equations of classical mechanics. Hopefully, however, this latter problem will be alleviated to some extent by the inclusion of occasional examples.

This Chapter is structured as follows. In the next section, I mention some of the difficulties surrounding infinite dimensional manifolds. This is the one bit of mathematical background for this Chapter which exceeds what is needed to read a textbook presentation of GR. I then turn to an exposition of the formalism of ordinary Hamiltonian mechanics from a geometric point of view. In §3 I discuss the symplectic geometry of classical phase spaces, and in §4 I discuss the way in which dynamics is modeled in this setting. §5 introduces presymplectic manifolds, which are slight generalizations of symplectic manifolds. §6 shows how such manifolds can be used to construct gauge theories, and §7 discusses how such gauge theories typically arise. §8 presents three strategies for dealing with the lack of uniqueness of solutions for the equations governing time evolution in gauge theories. Finally, §9 discusses how including time among the canonical variables of a Hamiltonian system transforms it into a gauge theory. This phenomenon is of considerable interest when trying to understand the role of time in classical and quantum gravity.

2. Infinite Dimensional Manifolds.

In the standard Hamiltonian formulation of classical point particle mechanics, the phase space of a system with N degrees of freedom is a 2N dimensional manifold, which represents the space of possible positions and momenta of the system. In field theories the objects under consideration typically have an infinite number of degrees of freedom. Thus we need to employ infinite dimensional manifolds to model the dynamical possibilities for such objects. This requires a generalization of the familiar theory of finite dimensional manifolds. This section provides a brief introduction to infinite dimensional manifolds, and the complications that accompany them.[46]

Recall that an *n dimensional manifold*, M, is a Hausdorff topological space which can be covered by a family of open neighborhoods, $\{U_\alpha\}$, such that each U_α is diffeomorphic to an open set of $V=\mathbf{R}^n$, and such that the diffeomorphisms mesh smoothly when two neighborhoods overlap. This amount of structure allows one to generalize most of the results of calculus of several variables. In particular, the local equivalence to V allows one to attach at each point $x\in M$ a copy of V, called the tangent space at x, or T_xM. T_xM allows one to talk about the tangent to a curve passing through x, or to look at the value at X(x) of a vector field, X, defined on M. Here a vector field is a smooth map $X{:}M\rightarrow TM$, where $TM=\cup_{x\in M}T_xM$ is the tangent bundle of M. Similarly, we can define T^*_xM to be the vector space dual to T_xM—so that elements of T^*_xM are just linear maps from T_xM to $\mathbf{R}$. We can go on to consider T*M, and covector fields, and higher order tensor spaces and tensor fields (by taking appropriate tensor products of the T_xM and T^*_xM). One can then employ the language of tensor algebra to formulate the differential and integral calculus on M .

What happens if we liberalize the definition of a manifold so that the vector space V to which M is required to be locally equivalent need not be finite dimensional, while holding fixed the other requirements? It turns out that we can carry over all of the tensor algebra and

[46] See Choquet-Bruhat, DeWitt-Morrette, and Dillard-Bleick 1982 for a detailed treatment, or Schmid 1987.

most of the calculus so long as we require V to be a Banach space (i.e., a normed vector space which includes the limits of all of its Cauchy sequences). In particular, we can prove versions of the Implicit Function Theorem, and the Chain Rule, which between them imply generalizations of many of the finite dimensional results. This motivates us to stipulate that an *infinite dimensional manifold* is a manifold modeled on an infinite dimensional Banach space.

Infinite dimensional manifolds do, of course differ in many significant ways from their finite dimensional counterparts. No infinite dimensional manifold is locally compact, for instance, although every finite dimensional manifold is. Furthermore, the fact that the tangent spaces are infinite dimensional leads to some complications which are not present in the finite dimensional case. Two are worth mentioning here. (i) In the finite dimensional case, a linear map $T:V\rightarrow V$ is one-to-one iff it is onto; in the infinite dimensional case a linear operator on V can be one-to-one but not onto. (ii) Similarly, in the finite dimensional case, all linear operators are continuous maps from V to V; in the infinite dimensional case, the continuous linear operators, i.e. the bounded operators, often have as their domain of definition a dense (proper) subset of V.

In what follows, I will attempt to be accurate, but will be willing to gloss over technicalities. Thus, I will be careful to characterize the Hamiltonian formalism so as to include the possibility of systems with infinitely many degrees of freedom. This will involve being careful about facts such as (i) above. But I will not attend to difficulties stemming from problems about domains of definition since in practice these can almost always be finessed by imposing some convenient restriction on the space of functions under consideration.

3. Symplectic Manifolds.

A *symplectic manifold* is a pair (M,ω), where M is a manifold and ω is a closed two-form on M which is non-degenerate in the sense that for each $x\in M$, the map

$_x$:T$_x$M→T*$_x$M defined by $_x$:v ω$_x$(v,·) is one-to-one.[47] M may be either finite or infinite dimensional; the conditions imposed on ω imply that if M is finite dimensional then dim M is even. (M,ω) is called *strongly symplectic* if each $_x$ is also onto. Every finite dimensional symplectic manifold is also strongly symplectic. If (M,ω) and (M',ω') are symplectic manifolds, then a diffeomorphism F:M→M' is called a *canonical transformation* if it preserves the symplectic structure (i.e., if F*ω'=ω).

In applications of symplectic geometry to classical mechanics, one usually thinks of M as being the space of possible dynamical states of some physical system, and of ω as providing M with some geometric structure. Typically, these dynamical states will just be the generalized positions and momenta of the system at a given time. I will call this interpretation of M the *standard interpretation*. It is not the only useful interpretation—we will see another below in §9. In thinking of ω as providing the manifold M with a geometric structure, it is helpful to keep in mind the analogy with the metric g of a semi-Riemannian geometry (N,g) (i.e., a manifold equipped with a non-degenerate, but not necessarily positive definite, metric). Both g and ω provide a bilinear map from V×V to **R**, where V is the tangent space at a point of their respective manifolds. Hence we can think of both as providing a sort of inner product, and thus as endowing their manifold with some geometric structure. Both, for instance, induce a notion of orthogonality (vw iff (v,w)=0). Furthermore, when (M,ω) is strongly symplectic, we can define ω^{-1}, the contravariant inverse to ω, and use it to raise indices just as in the semi-Riemannian case.

Of course, there are also important differences between symplectic geometry and semi-Riemannian geometry, stemming from the fact that ω is anti-symmetric while g is symmetric. Thus, the fact that in a semi-Riemannian geometry there are v∈T$_x$N such that g(v,v)>0 makes it possible to use to measure the length of a curve tx(t) in N. In the case of

---

[47] Recall that a k-form on M is an anti-symmetric tensor field of type (0,k). A form ω is closed if dω=0, where d is the exterior derivative. In components, $d\omega=\partial_a\omega$ if ω is a zero-form (function); $d\omega = 1/2(\partial_a\omega_b-\partial_b\omega_a)$ for one forms; $d\omega=1/6(\partial_a\omega_{bc}+\partial_b\omega_{ca}+\partial_c\omega_{ab}-\partial_b\omega_{ac}-\partial_c\omega_{ba}-\partial_a\omega_{cb})$ for two-forms, etc.

symplectic manifold, on the other hand, $\omega_x(v,w)=-\omega_x(w,v)$ for all $v,w\in T_xM$, so that $\omega_x(v,v)=0$ for all v. Thus ω does not endow M with any metric structure.

Another important disanalogy concerns the extent to which we can simplify our component representation of the respective geometric structures through careful choice of coordinates. In a generic semi-Riemannian geometry, (M,g), it is not possible to chose coordinates on a neighborhood U of $x\in M$ such that the components $g_{ij}(y)$ of the metric tensor are constant as y varies in U—although it is always possible to make the components constant on a given geodesic through x (there are, of course, exceptional, flat, semi-Riemannian geometries in which the components *can* be made locally constant). In a strongly symplectic manifold (M,ω), however, one can always choose *canonical coordinates* so that the components of ω are locally constant (this is a generalization of Darboux's theorem; it holds for some but not all infinite dimensional symplectic manifolds which fail to be strongly symplectic; see Marsden 1981, pp. 6-10). If (M,ω) is finite dimensional, this means that we can always choose local canonical coordinates $(p_1,\ldots p_n;q^1,\ldots,q^n)$ in which $\omega=dp_i\wedge dq^i$—so that as a matrix , where I is the $n\times n$ identity matrix. Thus any two symplectic geometries of the same finite dimension are locally isomorphic, something which is far from true in the semi-Reimannian case.

The chief utility of symplectic structure lies in the fact that ω induces a correspondence between functions on M and vector fields on M. If $f\in C^\infty(M)$, then the *Hamiltonian vector field* of f, X_f, is the solution to the equation $X_f\lrcorner\omega=df$.[48] There are couple of potential difficulties here. If M is infinite dimensional, then some perfectly good functions may not have a Hamiltonian vector field (this problem does not arise in the finite dimensional case). Even when X_f exists, its integral curves may be incomplete (i.e., the vector field is only locally defined). In the following, we will circumvent these problems by

[48] Here df is the exterior derivative of f—see the previous footnote. $X_f\lrcorner\omega:TM\rightarrow\mathbf{R}$ is the one-form $\omega(X_f,\cdot)$ whose value at $x\in M$ is $\omega_x((X_f)_x,\cdot)$. In components: $X_f^a\omega_{ab}$. This is known as the contraction of X_f with ω. When (M,ω) is strongly symplectic, we can write $X_f=df\lrcorner\omega^{-1}$.

restricting attention to the set $O(M,\omega)\subseteq C^\infty(M)$ of functions which have global Hamiltonian vector fields. The following operation, known as the *Poisson bracket*, gives $O(M,\omega)$ the algebraic structure of a Lie algebra: $\{f,g\}=\omega(X_f,X_g)$. It turns out that $[X_f,X_g]=X_{\{f,g\}}$, where $[\cdot,\cdot]$ is the Lie bracket, so that the map fX_f is a Lie algebra homomorphism of $O(M,\omega)$ into $\Xi(M)$, the algebra of vector fields on M.[49] Since we can also express the Poisson bracket as $\{f,g\}=X_f(g)$, we can think of $\{f,g\}$ as measuring the rate of change of g along the curves generated by f, so that g is constant along the curves generated by f iff $\{f,g\}=0$.

Each vector field $X\in\Xi(M)$ defines a *flow*, a one parameter family of diffeomorphisms ϕ_t:M→M, where $\phi^X{}_t(x)$ is the point of M which lies t units of time down the integral curve of X which passes through x. Thus we can think of each $f\in O(M,\omega)$ as the generator of the diffeomorphisms (x). The flow generated by f is canonical, in the sense that $\omega=*\omega$ for each t.

4. Hamiltonian Systems.

A *Hamiltonian system* is a triple (M,ω,H) where (M,ω) is a symplectic manifold, called the *phase space*, and H is a distinguished element of $O(M,\omega)$, called the *Hamiltonian*. This structure determines a unique preferred Hamiltonian vector field X_H, where $X_H\lrcorner\omega=dH$. The integral curves of X_H are called the *dynamical trajectories* of (M,ω,H). These curves, $t\in\mathbf{R}(p_i(t);q^i(t))$, are the solutions of *Hamilton's equations,* and . Every $x\in M$ lies on a unique such trajectory. In many physical applications, we follow the standard interpretation, and think of (M,ω) as being the space of generalized positions and momenta of some physical system, and H as being the energy. The dynamical trajectory through $x\in M$ describes the history of a system initially in the state represented by x. Notice that conservation of energy is a trivial consequence of the formalism—$\{H,H\}=\omega(X_H,X_H)=0$ so that H is a *constant of motion*, i.e. H is constant along each dynamical trajectory.

[49] The Lie bracket of $X,Y\in\Xi(M)$ is defined by $[X,Y](f)=X(Y(f))-Y(X(f))$, where $f\in C^\infty(M)$.

In general, one is interested in Hamiltonian systems in which the phase space is a cotangent bundle.[50] One starts with a manifold Q, the *configuration space* of the system. If one is doing particle mechanics, then each point in Q corresponds to a given spatial distribution of the particles relative to some fixed inertial frame; if one is doing field theory then each point corresponds to an instantaneous configuration of the field. One then looks at the cotangent bundle T*Q. Each point (q,p) in T*Q corresponds to a point in Q together with a covector, p, at that point. One interprets p as representing the momentum of the configuration q. Every cotangent bundle T*Q has a canonical symplectic structure ω, so that as soon as one specifies a Hamiltonian, one has a theory in which specification of the position and momentum of the system at a given time determines its dynamical history.

The procedure for finding this canonical structure is straightforward in the finite dimensional case (it is only slightly more complicated in the infinite dimensional case; see Schmid 1987, p. 51). One first chooses a set of local coordinates $(q^1,\ldots,q^n)$ on Q. Then $\{dq^i\}$ is a basis on each T^*_qQ. Writing each $v \in T^*_qQ$ as $v=p_i dq^i$ gives us coordinates $(q^1,\ldots,q^n;p^1,\ldots,p^n)$ on T*Q. We can then stipulate that these are canonical coordinates, and write $\omega=dp_i \wedge dq^i$. In this case, we call p_i the *momentum canonically conjugate* to q^i. This construction of ω is independent of the original choice of coordinates on Q—that is, changes of coordinates on Q generate canonical transformations.

Examples:

(1) If (M,ω,H) is a finite dimensional Hamiltonian system and $(p_1,\ldots,p_n;q^1,\ldots,q^n)$ are canonical coordinates, then the dynamical trajectories are just the solutions of Hamilton's familiar equations and or, equivalently,

and .

[50] The most important exceptions are systems of classical point particles with spin. There the internal degrees of freedom appear as momentum variables which do not correspond to any position variables, so that the phase space is not naturally isomorphic to a cotangent bundle (see §7.1 of Woodhouse 1980).

(2) If we are dealing with a free particle in Euclidean space, then Q=$\mathbf{R}^3$. If our position variable is just the position of the particle relative to some Cartesian system of coordinates, then the canonically conjugate momentum is the ordinary linear momentum, and H=. More generally, if we have a free particle on a Riemannian manifold (M,g) then Q=M and the phase space is T*M. Then $H=1/2(g^{ab}p_a p_b)$ generates the dynamical trajectories which correspond to the particle moving along the geodesics of (M,g) (here the p's are the momenta canonically conjugate to an arbitrary set of coordinates on Q).

(3) Fix an inertial frame in Minkowski space, and let Q be the space of configurations relative to this frame of Klein-Gordon field ϕ of mass m—thus each point in Q corresponds to a $\phi:\mathbf{R}^3\rightarrow\mathbf{R}$. We then look at T*Q, where a point corresponds to a pair . The symplectic form is given by

where Σ is a time-slice of our inertial frame (here we are dealing with a linear field theory, so that Q and T*Q are vector spaces, and we may identify vectors on T*Q with elements of T*Q; see Wald 1994). The Hamiltonian is

H=.

Both ω and H are independent of the choice of Σ. The equation of motion is just the usual Klein-Gordon equation, .

5. Presymplectic Manifolds.

A *presymplectic manifold* is a pair (N,σ) where N is a manifold and σ is a degenerate, closed two-form on N. That is, there is an $x\in N$ such that the map $_x:T_xN\rightarrow T^*_xN$ defined by $_x:v\sigma_x(v,\cdot)$ is not one-to-one. Thus ker $_x=\{v\in T_xN: \sigma_x(v,\omega)=0$ for all $w\in T_xN\}$ is non-trivial. For convenience, we will assume that dim ker $_x$ is constant as x varies in N (see Appendix B.3 of De León and Rodrigues 1985 for a sketch of the general theory).

Thus the only difference between a symplectic manifold and a presymplectic manifold is that a presymplectic form has null vectors (i.e., the elements of ker). So presymplectic manifolds are in some ways quite similar to symplectic manifolds. Most importantly: given a presymplectic (N,σ) and a f:N→**R**, it makes sense to write down and solve the equation $X_f \lrcorner \sigma = df$; the solutions are again known as the Hamiltonian vector fields of f. Again, we restrict attention to the class of functions which have global Hamiltonian vector fields, O(N,σ). But, whereas in the symplectic case the Hamiltonian vector field of f was unique (if it existed), in the presymplectic case there will be many Hamiltonian vector fields. Indeed let X and X' be vector fields on N, and suppose that X is a Hamiltonian vector field for f. Then X' is a Hamiltonian vector field for f iff Y=X–X' is a vector field whose value at each point x∈N is a null vector of σ_x (since $0=X\lrcorner\sigma - X'\lrcorner\sigma = Y\lrcorner\sigma$).

Despite the failure of uniqueness for Hamiltonian vector fields, we can still define the Poisson bracket of f,g∈O(N,σ) as $\{f,g\}=\sigma(X_f,X_g)$ (since if X, X', and Y are as above, then $\sigma(X,X_g)-\sigma(X'X_g)=\sigma(X'+Y,X_g)-\sigma(X'X_g)=\sigma(Y,X_g)=0$). We can still think of {f,g} as measuring the rate of change of g along the direction determined by f as an infinitesimal generator—except that f now generates an entire family of curves. Thus {f,g}=0 iff f is constant along the flows of *all* of the Hamiltonian vector fields of g.

6. Gauge Systems.

A *gauge system* is a triple (N,σ,H) where (N,σ) is a presymplectic manifold, and H is a distinguished element of GO(N,σ). As in the Hamiltonian case: (N,σ) is called the phase space, H is called the Hamiltonian, and the integral curves of each X_H are called dynamical trajectories (X_H is any solution of $X_H \lrcorner \sigma = dH$). There are infinitely many dynamical trajectories through each x∈N—although, of course, any two Hamiltonian vector fields, X and X', of H differ by a null vector field Y.

In applying this formalism, it turns out to be interesting to divide N into equivalence classes as follows: for x,y∈N, we say that x~y if x and y can be joined by a curve, x:

$t\in[a,b]x(t)\in N$, whose tangent vectors are all null vectors of σ (i.e., for all $t\in[a,b]$). Then it follows from Frobenius' theorem and the fact that $[X_f,X_g]=X_{\{f,g\}}$ that $[x]=\{y\in N: x\sim y\}$ is a submanifold of N with dim [x]=dim ker .[51] We call [x] the *gauge orbit* of x, and call a diffeomorphism $F:N\rightarrow N$ a *gauge transformation* if it preserves gauge orbits (i.e. if $F(x)\in[x]$ for all $x\in N$). Thus a null vector field Y on (N,σ) can be regarded as an infinitesimal generator of a gauge transformation, since the flow generated by Y consists of gauge transformations. In what follows, we will often be interested in the set $GO(N,\sigma)=\{f\in O(N,\sigma)$: f is constant of each $[x]\}$ of *gauge-invariant* elements of $O(N,\sigma)$.

7. Constrained Hamiltonian Systems.

We will see in Chapter 7 that GR is a gauge system. So are the classical limits of the quantum theories of the electromagnetic, strong nuclear, and weak nuclear interactions. Thus much of contemporary mathematical physics is given over to the study of gauge theories.[52] Typically these theories arise in the following way. One starts off with a Lagrangian formulation of some theory. That is, one represents the set of states of the system using the tangent bundle, TQ, of some configuration space Q, and specifies a smooth function $L:TQ\rightarrow\mathbf{R}$, the *Lagrangian*, which determines the dynamics via Hamilton's Principle.[53] There is a canonical technique for passing from this formalism to the Hamiltonian formalism, called the *Legendre transform*. If (where $(q^1,\ldots q^n;v^1,\ldots v^n)$ are the obvious coordinates on TQ), then the Legendre transform gives us a diffeomorphism between TQ and T*Q. One can then go on to define the so-called canonical Hamiltonian, and observe

[51] See §2.2.2 of Henneaux and Teitelboim 1992 for a proof in the finite dimensional case, and Marsden 1981, p. 6 for a sketch of the infinite dimensional treatment. See §IV.8 of Boothby 1986 or §IV.6 of Choquet-Bruhat, DeWitt-Morrette, and Dillard-Bleick 1982 for Frobenius' theorem.

[52] It is somewhat more difficult to find physically realistic gauge theories with finitely many degrees of freedom. Here the theories of Barbour and Bertotti 1982 and Lynden-Bell 1995 are probably the most interesting candidates. These are discussed below. See also §3 of Kuchar 1992 for a number of examples which, although not physically realistic, are nonetheless thought to provide illuminating analogies to quantum gravity.

[53] Hamilton's principle states that a given curve $\gamma:[t_1,t_2]\rightarrow Q$ of the form $tq^a(t)$ corresponds to a dynamical trajectory iff $I(\gamma)=$ is stable under variations of γ where we fix the endpoints.

that the images under the Legendre transform of the dynamical trajectories are the dynamical trajectories of the canonical Hamiltonian on T*Q. If, however, , then the Legendre transform is not a diffeomorphism because it is not one-to-one: the image of TQ, N, is a proper submanifold of T*Q. Thus, if L is singular the Lagrangian theory does not correspond to a Hamiltonian system, but rather to a submanifold N, of a cotangent bundle, together with a function $H:N\rightarrow\mathbf{R}$. If we look at the structure that the canonical symplectic structure on T*Q, ω, induces on N as a submanifold of M, then we find that $\sigma=\omega|_N$ is typically a presymplectic form on N, so that (N,σ,H) is a gauge system.[54]

This motivates the following definition: a *constrained Hamiltonian system* is a gauge theory (N,σ,H) where (N,σ) is a regular submanifold of a symplectic manifold (M,ω) (i.e., $N\subset M$ and $\sigma=\omega|_N$). Since we will often be interested in the restriction to N of functions on M, we introduce the notation $f\cong g$ (read "f is weakly equal to g") for $f,g\in O(M,\omega)$ such that $f|_N=g|_N$.

At least locally, we can always specify N by requiring that some set, $C=\{c_a\}_{a\in I}$ of real-valued functions on M vanish. These functions are called *constraints*.[55] There are two kinds of constraints: if $c_a\in C$ is such that the restriction to N of the Hamiltonian vector field of c_a is a null vector field of (N,σ), then c_a is a *first class constraint*, and is denoted γ_a; otherwise, c_a is a *second class constraint*, and is denoted χ_a. Thus the first class constraints, but not the second class constraints, generate gauge transformations on N. In fact, at each point $x\in N$ {} is a basis for ker , so that dim ker =card $\{\gamma_a\}$. This means that if $f\in O(M,\omega)$, then $f|_N\in GO(N,\sigma)$ iff $\{f,\gamma\}\cong 0$ for all first class constraints γ (here {,} are the Poisson

[54] Here an analogy with Minkowski spacetime may be helpful. If we look at the metric on a two dimensional submanifold of Minkowski spacetime, then we often find that the metric induced on this submanifold by the Minkowski metric is degenerate in the sense that it has null vectors. It can happen that the dynamics induced by H carries points off the submanifold which is the image of the Legendre transformation. In this case, one restricts attention to a smaller submanifold, which is the end result of the Dirac-Bergmann algorithm (see §1.1.5 of Henneaux and Teitelboim 1992).

[55] One has to be a bit careful here: not just any set of functions will do. One requires that the constraint functions be a subset of some set of local coordinates on M; see §1.1.2 of Henneaux and Teitelboim 1992. This is always possible since N is by assumption a regular submanifold of M.

brackets on (M,ω)). One describes this result by saying that the elements of $GO(N,\sigma)$ *commute with the first class constraints.*

Having imbedded (N,σ,H) in (M,ω), it is natural to wonder about the relationship between the dynamical trajectories of (N,σ,H) and the restriction to N of the Hamiltonian vector fields of (M,ω). We call $h\in O(M,\sigma)$ an *extension* to (M,ω) of H if: (i) $h|_N=H$; (ii) $\{h,c\}\cong 0$ for all constraints c. The latter condition means that flow generated by h carries points on N to points on N, since the Hamiltonian vector field of h is everywhere tangent to N. If h is an extension of H, and X_h is the Hamiltonian vector field of h in (M,ω), then $X_h|_N$ is a Hamiltonian vector field of H in (N,σ). Conversely, every Hamiltonian vector field of H in (N,σ) arises in this manner, for some extension h of H. It is not difficult to prove that any two extensions, h and h' of H differ by a linear combination of first class constraints. It follows that the transformation $h\rightarrow h+u^a\gamma_a$ carries us from one set of dynamical trajectories of (N,σ,H) to another, where the u^a are arbitrary functions on M; conversely, every pair of sets of dynamical trajectories are so related (see §§1.1 and 1.2 of Henneaux and Teitelboim 1992 or Chapter 1 of Dirac 1964).

Whereas in the Hamiltonian case, Hamilton's equations and determine a unique dynamical trajectory of the form $t\in\mathbf{R}(p_i(t);q^i(t))$ through each $x\in M$, we see that in the case of a constrained Hamiltonian system, Hamilton's equations and determine a different set of dynamical trajectories for each h which extends H. Given our freedom to replace h by $h'=h+u^a\gamma_a$, we can write Hamilton's equations as and . Thus the solutions of Hamilton's equations, which determine the dynamical trajectories of (N,σ,H) contain as many arbitrary functions of time as there are first class constraints.

Examples:

(4) Let (M,ω) be a symplectic manifold with canonical coordinates $(p_i;q^i)$ and suppose that (N,σ,H) is a constrained Hamiltonian system with N given by the first class constraint $p^1=0$. Then the Hamiltonian vector field of p_1 in (M,ω) generates motions in the q^1 direction. The

gauge orbits in (N,σ,H) are of the form $\{(0,p_2,\ldots;s,q^2,\ldots): s\in\mathbf{R}$, all other p_i and q^i fixed$\}$. So $GO(N,\sigma)$ is the subset of $O(N,\sigma)$ which consists of functions which are independent of q^1. A dynamical trajectory is of the form $t\in\mathbf{R}(p_i(t);q^i(t))$; Hamilton's equations of motion determine $p_i(t)$ and $q^i(t)$ uniquely for $i\geq 2$, while allowing $q^1(t)$ to be an arbitrary function of time, and leaving $p_1(t)=0$. The behavior of this trivial example is typical: if (N,σ,H) is a constrained Hamiltonian system in a finite dimensional symplectic manifold (M,ω), then we can always find canonical coordinates, $(p_i;q^i)$, on (M,ω) so that the first class constraints are of the form $p_i=0$ for $i\leq k$, so that the $q^i(t)$ is arbitrary for $i\leq k$ (see Theorem 5.1.11 of Abraham and Marsden 1978).

(5) Lynden-Bell's first theory (LB1).[56] This is a theory of the gravitational interaction of n particles in Euclidean space. Let $Q=\mathbf{R}^{3n}=\{\mathbf{x}^i=(x^i,y^i,z^i)\in\mathbf{R}^3: 1\leq i\leq n\}$ be the coordinates of the particles relative to some Cartesian coordinate system on $\mathbf{R}^3$. Our phase space is T^*Q with the symplectic structure ω induced by stipulating that the momentum canonically conjugate to $\mathbf{x}^i$ is:

,

where m_i is the mass of the ith particle. We then look at the submanifold $N\subset T^*Q$ given by constraint , with the induced presymplectic form $\sigma=\omega|_N$ and the Hamiltonian

.

(Notice that the second term in the expression for H is just the ordinary potential energy of Newtonian Gravity.) The gauge orbits of (N,σ,H) are constructed using the equivalence relation $(\mathbf{p}_i;\mathbf{x}^i)\sim(\mathbf{p}_i';\mathbf{x}^{i\prime})$ iff $\mathbf{p}_i=\mathbf{p}_j'$ and $\mathbf{x}^j-\mathbf{x}^j=\mathbf{a}$, for all j, with $\mathbf{a}\in\mathbf{R}^3$ independent of j. Thus the gauge transformations are just spatial translations of the system of particles in Euclidean space. In terms of the equations of motion, this means that any two dynamical trajectories through a given point in N will differ in their representation of the history of the particles in

[56] See Lynden-Bell 1995 for this example and the next. Thanks to Carlo Rovelli, for translating Lynden-Bell's Lagrangian formulation into the Hamiltonian language.

$\mathbf{R}^3$ by an arbitrary time-dependent boost $\mathbf{a}(t)$ (thus there are three arbitrary functions of time in the solutions to the equations of motion).

This theory gives the same relative distances between particles as the standard Newtonian theory. The difference lies in the roles that inertial frames play in the two theories: in both theories there is an absolute standard of rotation (i.e., neither theory is invariant under time dependent rotations); in the Newtonian theory the structure of spacetime also provides an absolute standard of linear acceleration; no such standard is available in LB1 (since if we fix a point in $x \in N$, and two dynamical trajectories through x, we find that in general if the histories of the particles described by the two trajectories are imbedded into Newtonian spacetime, they will in general disagree concerning the absolute acceleration of the particles). For this reason it is tempting to see LB1 as the result of incorporating the Equivalence Principle into Newtonian gravity (compare with the analysis of the Newton-Cartan theory in Norton 1995).

(6) Lynden-Bell's second theory (LB2). Again, this is a theory of n gravitating point particles, with the same configuration space Q as in example 5. Let $\mathbf{p}_i$ be the momentum canonically conjugate to $\mathbf{x}^i$, and let T*Q have the standard symplectic structure ω (I omit the somewhat messy details). This time N is given by the constraints and . After imposing an appropriate Hamiltonian H on $(N,\omega|_N)$, we get a theory of gravity which is equivalent to the Newtonian one iff the center of mass of the universe has zero angular momentum. The gauge orbits of the system are the equivalence classes given by $(\mathbf{p}_i;\mathbf{x}^i)\sim(\mathbf{p}_i';\mathbf{x}^{i'})$ iff $\mathbf{p}_i=\mathbf{p}_j'$ and $\mathbf{x}^j=O\mathbf{x}^j+\mathbf{a}$, with $\mathbf{a}\in\mathbf{R}^3$ and O a rotation on $\mathbf{R}^3$ both independent of j. That is, $(\mathbf{p}_i;\mathbf{x}^i)\sim(\mathbf{p}_i';\mathbf{x}^{i'})$ they differ by the action of an orientation preserving Euclidean isometry. In terms of the equations of motion, two solutions by a time dependent rotation as well as a time dependent acceleration. Thus, as in the case of LB1, there is not standard of absolute linear acceleration. Furthermore, there is no standard of absolute rotation—this is wiped out by time dependent rotations.

(7) Vacuum Electromagnetism. Let (S,g) be a three dimensional manifold representing physical space.[57] And let Q={A: S→$\mathbf{R}^3$} be the space of covector fields on M—that is, each element of Q is a function which maps each point of S to a three dimensional vector. We call such a map, A, a vector potential—although no particular meaning should be attached to that name for the time being. We construct the cotangent bundle, T*Q, and endow it with the canonical symplectic structure, ω. A point in T*Q is a pair (E,A), where E, like A, is a vector field on S.

In order to construct the phase space of electromagnetism, we impose the constraint div E=0. This gives us an infinite dimensional submanifold, N, of T*Q.[58] The restriction to N of the canonical symplectic structure on T*Q gives us a presymplectic form, σ. Thus, (N,σ) is a presymplectic geometry. This is the phase space of electromagnetism. The gauge orbits of (N,σ) are determined by the following equivalence relation: (E,A)~(E',A') iff E=E' and A'=A+gradΛ for some Λ: S→$\mathbf{R}$. Thus, [(E,A)]={(E,A'): A'=A+gradΛ for some Λ: S→$\mathbf{R}$}.

We now specify the Hamiltonian H=. Hamilton's equations become:

.

---

[57] The metric g plays a hidden role in what follows. In ordinary Euclidean space, we move back and forth between vectors and covectors, and between one forms and two forms without noticing. In a general three dimensional geometry, one has to be more careful. Thus, whereas in three dimensional Euclidean space, we can write grad f=∇f, div v=$\nabla\cdot$v, and curl v=$\nabla\times$v, we must exercise caution in a general three dimensional geometry.

Let J be the isomorphism which g induces between vectors and covectors. And let $*$, the *Hodge star*, be the isomorphism which g induces between one forms and two forms (and its inverse). This latter is determined by the map which maps the one form, $B=B_x dx+B_y dy+B_z dz$, to the two form

$*$B=,

where the vector (B^x,B^y,B^z) is $g^{ab}B_a$. We now define: grad f=J^{-1}df, div v=-$*$d$*$Jv, and curl v=$J^{-1}*$dJv.

[58] N, Q, and T*Q are infinite dimensional vector spaces. Thus they are simply connected—no matter what the topology of S. This is in sharp contrast to the case of a phase space of a single particle moving on S, which is just T*S, and hence is simply connected iff S is. We will see below, however, that the structure of N *does* depend in an interesting way on the topology of S.

These are Maxwell's equations for E and A, the electric field and the vector potential. Here we find the behavior that we expect from a gauge system: specifying an initial point does not serve to determine a unique dynamical trajectory. But we do find that if (E(t),A(t)) and (E'(t),A'(t)) are solutions of (∗) for the initial data $(E_0,A_0)\in N$, then for each t E(t)=E'(t) and there is a scalar function on space, Λ(t), such that A'(t)=A(t)+grad Λ(t). Equivalently: if (E(t),A(t)) and (E'(t),A'(t)) are dynamical trajectories through the same point of (N,σ), then we have that [(E(t),A(t)]=[(E'(t),A'(t)] for all t. Maxwell's equations do not determine the future value of A(t), but they do determine which gauge orbit A(t) will lie in.

8. Three Strategies for Coping with Gauge Freedom.

In order to use these abstract formalisms to model classical physical systems, we will want to be able to write down differential equations, plug in initial values for the variables, and crank out unique solutions which describe the evolution of a system in the given initial state. That is: we expect that if f is a function on the space of states which corresponds to a physical quantity, then the *initial value problem for f is well-posed*, in the sense that specifying the initial data for the differential equations governing the evolution of f serves to fix a unique solution. It is very difficult to do classical physics in the absence of a well-posed initial value problem (see Earman 1986 and Wilson 1989 for examples and discussion).

The situation is straightforward when we are dealing with a Hamiltonian system (M,ω,H). If $f\in O(M,\omega)$, we can take as our initial data the coordinates of $x\in M$. The initial value problem for f is then well posed, since the fact that there is a unique dynamical trajectory of (M,ω,H) which passes through x means that there exists a unique solution of Hamilton's equation .

In the case of a gauge system (N,σ,H), on the other hand, it is easy to see that the initial value problem for position and momentum is not well-posed, since there are infinitely many dynamical trajectories though each point of N. The presymplectic structure is simply inadequate to single out a unique solution for a given set of initial data. However, as we have seen above, the gauge freedom to move from one set of dynamical trajectories to another is constrained in an interesting way by the presymplectic structure: two Hamiltonian vector fields X and X' of H must differ by a null vector field; if N is a constraint surface in (M,ω), this means that X and X' are restrictions to N of Hamiltonian vector fields in (M,ω) whose generators differ by a linear combination of first class constraints. One might hope that it would possible to exploit these relations to render gauge systems useful for modeling classical physical systems. This is indeed the case. There are three closely related standard strategies for managing the gauge freedom intrinsic to a gauge system (N,σ,H).

(i) Gauge Invariant Functions.

Fix $x \in N$. If tx(t) and tx'(t) are two dynamical trajectories which pass through x at t=0, then in general we expect that $x(t) \neq x'(t)$ for t>0 (the initial value problem for position and momentum is ill posed). But the presymplectic structure ensures that [x(t)]=[x'(t)] for all t. That is, although the various dynamical trajectories passing through x disagree about state of the system is in at time t, they agree about which gauge orbit the system is in at time t. This means that the initial value problem for $f \in O(N,\sigma)$ is well-posed iff f is an element of GO(N,σ), the set of gauge invariant functions on N.

If N is a constraint surface in (M,ω), then $f \in O(M,\omega)$ is also called gauge invariant if $f|_N \in GO(N,\sigma)$. i.e., if $\{f,\gamma\} \cong 0$ for every first class constraint γ. Let $h \in O(M,\omega)$ be an extension of H to (M,ω). Then gauge transformations $h \to h' = h + u^a\gamma_a$ of h generate all and only the Hamiltonian vector fields of H when restricted to N. If $f \in O(M,\omega)$, then evolution of f along the dynamical trajectories of (M,ω,h) is given by . If f is gauge invariant we have:

$$\{f,h'\} \cong \{f,h+u^a\gamma_a\} \cong \{f,h\}+\{f,u^a\gamma_a\} \cong \{f,h\}.$$

So the equations governing the evolution of on N are independent of the choice of h on (M,ω). Which is just to say that the evolution of the gauge invariant functions is uniquely determined by the structure (N,σ,H).

In example (4), $f \in GO(N,\sigma)$ iff it is independent of the value of q^1, so that we simply ignore the gauge degree of freedom. In example (5), relative distances and absolute accelerations are gauge invariant, but absolute linear acceleration is not. In example (6), absolute rotations is not gauge invariant—in fact, all gauge invariant functions are now functions of relative distances between particles. In example (7), the gauge invariant functions on T*Q are those which are invariant under the gauge transformations $A \to A' = A +$. When space is simply connected, these turn out to be the functions which can be re-written as functions of E and and the magnetic field Bcurl A (the situation is more complicated if space is multiply connected; see §6.8 below).

(ii) Gauge Fixing.

Our first strategy for finding a well posed initial value problem involved restricting our attention to a subset of $O(N,\sigma)$. Our second strategy involves restricting attention to a subset of N. Our idea here is to exploit the fact that the dynamical trajectories agree about the sequence of gauge orbits which the system passes through, by restricting our attention to a submanifold N' of N which contains one representative of each gauge orbit. This is known as *fixing the gauge*. If all works out well, then (N',σ',H'), where $\sigma' = \sigma|_{N'}$ and $H' = H|_{N'}$, is a Hamiltonian system and every $f' \in O(N',\sigma')$ has a well posed initial value problem.

This is easiest to visualize in the case where is N the submanifold in (M,ω) which satisfies the constraints $\{C_a\}_{a \in I}$. Here we proceed by adding further constraints $\{D_b\}_{b \in J}$ in such a way that the new constraint manifold $N' \subset N$ intersects each gauge orbit of (N,σ,H) exactly once. This means that $\sigma' = \sigma|_{N'}$ has no null vectors (since the null vectors of σ generate gauge motions along N which are not tangent to N'). Thus the new constraint set $\{C_a\} \cup \{D_b\}$ consists entirely of second class constraints, and (N',σ',H') is a Hamiltonian system.

In example (4), this can be accomplished by requiring that $q^1 = k$, for some $k \in \mathbf{R}$. By arbitrarily fixing the gauge degree of freedom we have transformed the gauge system into a

Hamiltonian system in which the behavior of p_i and q^i for $i \geq 2$ is deterministic. In examples (5) (and (6)), the gauge can be fixed by stipulating that the center of mass of the universe is at rest (and non-rotating). We then recover the ordinary Newtonian formulation. In electromagnetism we can impose the Lorentz gauge condition $\partial^a A_a = 0$ as an additional constraint. Maxwell's equations then have a well-posed initial value problem, and we have succeeded in eliminating the gauge freedom $A \rightarrow A' = A +$.

There are three points worth noting. (1) Gauge fixation is not always possible. In particular, if N has a non-trivial topology, it is often impossible to find a submanifold N' which intersects each gauge orbit exactly once (this is known as the Gribov obstruction; see §1.4.1 of Henneaux and Teitelboim 1992). (2) Any Hamiltonian system (M,ω,H) can always be imbedded in a gauge system (N,σ,H') in such a way that the Hamiltonian system arises from the gauge system by gauge fixation. (3) If we have a constrained Hamiltonian system, then every gauge invariant function on the constraint surface has many gauge invariant extensions to the embedding space. But if we have a function on a gauge-fixed constraint surface, then it has a unique gauge invariant extension to the phase space of original gauge system.

(iii) Reduced Phase Space.

Our goal is to isolate the well posed initial value problem at the core of the gauge system (N,σ,H). The first strategy does so by restricting attention to a subset of the set of functions. This allows us to achieve our goal, but leaves us with a great deal of redundant structure in our phase space, since we ignore the internal structure of each gauge orbit. The second strategy discards this excess structure, by restricting attention to a subset of the phase space. Again, this leaves us with a well posed initial value problem, but at the price of having introduced an apparently arbitrary restriction. Throwing away most of the original states is likely to seem unattractive, unless one believes, as Feynman is reputed to have done, that the additional constraints that serve to fix the gauge have real physical content (see §6.7 for further discussion of this point).

Our third strategy seeks to improve on its predecessors by constructing a new phase space, whose points are the gauge orbits of the original gauge system. That is, one attempts to construct a symplectic geometry, the *reduced phase* space (M',ω'), where M' is the manifold N/~ and ω' is the symplectic form on M' inherited from (N,σ). The points of M' are the gauge orbits, [x], of (N,σ) with the strongest topology which makes the projection map, π: x[x], continuous, while $\sigma=\pi^*\omega'$.The Hamiltonian H on (N,σ) induces a unique Hamiltonian h on (M',ω'), so that all of the Hamiltonian vector fields of H project down to the single Hamiltonian vector field of h. Thus we again achieve a well-posed initial value problem, this time without making any apparently arbitrary restrictions.

In example (4), the reduced phase space is just $\{(p_2,\ldots p_n;q^2,\ldots q^n)\}$ with the obvious symplectic structure and Hamiltonian. In example (5), the reduced phase space arises when points in T*Q which differ by a spatial translation are identified. In example (6), the phase space is constructed by identifying points which differ by an isometry. As discussed in §2.12, this means that the reduced phase space is just T^*Q_0, where Q_0 is the space of relative distances rather than positions relative to some coordinate system. In the case of electromagnetism on a simply connected physical space, the reduced phase space is just the set {(E,B): div E=div B=0} with the Hamiltonian (the situation is more subtle when spacexq has a more complicated topology; see §6.8).

Despite its elegance, the reduced phase space approach has some drawbacks. First of all, one can lose desirable features in passing to the reduced phase space, such as a natural cotangent bundle structure or manifest Lorentz covariance (see Appendix 3 of Ashtekar and Geroch 1974, and §2.2.3 of Henneaux and Teitelboim 1992, respectively).[59]

---

[59] Either of these deficiencies can cause trouble when it comes time to quantize the theory. (i) If a theory is Lorentz covariant, but not manifestly Lorentz covariant, it is not a foregone conclusion that its quantization will be Lorentz covariant. (ii) If the phase space of a theory can be written as T*Q in a natural way, then one can go on to construct the Schrödinger representation of its quantization by taking $L^2(Q)$ as the Hilbert space. Quantization is still

Furthermore, of course, the set of gauge orbits need not be a manifold, so that it may be impossible to construct the reduced phase space (this is distinct from the Gribov obstruction; see Appendix 2.A of Henneaux and Teitelboim 1992). It occasionally happens, for instance, that M' fails to be Hausdorff, so that it cannot be a manifold (see §2.6 of Woodhouse 1980 for a simple example). Or M' can be Hausdorff, but still fail to be a manifold in virtue of having conical singularities. This happens, for instance, in the case of the space of globally hyperbolic models of GR with compact Cauchy surfaces.[60]

9. Parameterized Systems.

One treats time and space very differently in the standard approach to Hamiltonian systems. The points in phase space represent the possible positions and momenta of the system relative to some fixed inertial frame, and the dynamical trajectories represent the physically possible histories of the system. Time, unlike space, is not directly represented by means of a canonical variable—rather, it stands outside of the system of variables and governs their behavior via the time derivatives in Hamilton's equations. It *is* possible, however, to include time among the canonical variables, and thus to treat space and time more symmetrically. This is known as parameterization. It is worthwhile to review this construction here, both because it will serve to drive home the point that there are a number of interpretations of a given formalism and because GR is in many ways similar to a parameterization of a Hamiltonian systems.

Let (M,ω,h) be a Hamiltonian system. We construct $M'=\mathbf{R}^2\times M$ by adding to M the canonically conjugate variables t and u. Let σ be the symplectic form on M' given by $\omega'=\omega-dudt$. Let $H=h+u$ and let N be the submanifold of M' determined by the constraint $H=0$ (we extend h to M' in the obvious way, by making it independent of t and u). Then the constrained Hamiltonian system (N,σ,H), with $\sigma=\omega'|_N$ and the Hamiltonian given by $H=0$,

possible when there is no such structure, but it often requires more resourcefulness; see Woodhouse 1980.

[60] See Lecture 10 of Marsden 1981. This problem does not arise in the space of asymptotically flat solutions; see §3.3 of Ashtekar, Bombelli, and Reula 1991.

is called the *parameterization* of (M,ω,h). We can think of (N,σ,H) as the result of including time among the position variables of the system, with the energy h as its canonically conjugate momentum (since h=–u on N). Notice that (N,σ) is presymplectic (in the finite dimensional case this is obvious since dim N is odd), with dim ker =1 everywhere on N. The fact that H0 means that the solutions of $X_H \lrcorner \sigma = dH = 0$ are just the null vector fields of σ.

Each dynamical trajectory on (M,ω,h) corresponds to a gauge orbit on (N,σ,H). Pick a time t and a point x∈M, and look at the dynamical trajectory t(p(t),q(t)) on M. Then the dynamical trajectories through (t,x)∈N will be of the form t(τ, (p(τ),q(τ))), where τ(t) is some re-parameterization of time. The gauge orbit in N which corresponds to the trajectory t(p(t),q(t)) in M will include all the points in N which are images of the maps t(τ, (p(τ),q(τ))), for all parameters τ. Thus the loss of the preferred parameterization of time is the price of including time among the canonical variables.

Of course, we can use any of the three strategies from the previous section to help us cope with the gauge freedom brought on by parameterization. Thus one can work with (N,σ,H') and simply restrict attention to gauge invariant functions. Here these are just the constants of motion—since time evolution is a gauge transformation, the only gauge invariant quantities are those which are time-independent. Or it is possible to pass to the reduced phase space, each of whose points corresponds to a gauge orbit of (N,σ,H)—or, equivalently, to a dynamical trajectory of (M,ω,h). Hamiltonian evolution on the reduced phase space amounts to translation of the origin of the time axis for the dynamical trajectories in the original Hamiltonian system. This is the promised example of a Hamiltonian system where points in M do *not* correspond to the possible positions and momenta of some physical system.

In general, however, it is most interesting to fix the gauge. Here the idea is to restrict attention to the submanifold N' of N which corresponds to dynamical trajectories which are 'properly parameterized' (i.e., a dynamical trajectory τ(p(τ),q(τ),u(τ),t(τ)) is on N' iff =1).

This process of gauge fixation is known as *deparameterization*. This allows us to recover the original Hamiltonian system.

We will see in Chapter 8 that each of these strategies plays an important role in the discussion of the nature of time in Quantum Gravity.

Example:

(8) The theory of Barbour and Bertotti 1982 (BB). This is equivalent to LB2, but is not just its parameterization (since the constraint governing time evolution is quadratic in the momenta). The proper choice of time function does, however, recast the theory as the parameterization of LB2.

CHAPTER 6
Determinism and Ontology: Interpreting Gauge Theories

1. This Chapter provides a discussion of some of the interpretative problems surrounding gauge theories. Here I have two main goals. The first, which is the object of §§2-4, is to make a general point about the connection between determinism and ontology. A popular view concerning determinism is that it is a *formal* property, independent of interpretation. This view, if correct, would provide a very quick refutation of the hole argument—since the latter purports to draw a connection between and indeterminism and an interpretative strategy (substantivalism), it *must* be fallacious. I argue that this formal view of determinism is untenable. My argument is driven by examples of gauge theories which admit multiple interpretations, under some of which they are deterministic, under some of which they are indeterministic.

My second goal, pursued from §5 onwards, is to develop a vocabulary and a conceptual framework in which to address the problem of finding an appropriate ontology for interpreting gauge theories. Here I sketch a number of interpretative options, and enumerate their merits and demerits. Once again, I rely on examples to illustrate my point. This discussion provides the foundation on which I will base my analysis of the interpretative problems of classical and quantum gravity in the following Chapters.

Before embarking on this course, I would like to say a few words about the notion of interpretation that I will be utilizing. Here I follow the logical empiricist tradition in using the word 'theory' to denote a purely linguistic object. A theory will be some fragment of the natural language of mathematics. As such it will be partially interpreted—the symbols will always be given their intended mathematical interpretation. But it will not be given a physical interpretation. Thus, a theory will typically be a set of differential equations. Maxwell's equations—together with the space of functions on which they are defined and the space upon which the functions occurring in the equations are defined—count as a theory. This is a *very* thin notion of theory—it does not include an intended physical interpretation, an associated experimental practice, or any relation to data.

With my notion of (physical) interpretation, I part ways with the logical empiricists. By an interpretation of a theory, I mean an account of the way in which the theory represents a set of possible worlds. This is an open-ended conception of interpretation. Most obviously, an interpretation includes an account of which structures in the formalism of the theory correspond to which bits of the possible worlds. A second important component is the measurement theory, which tells us how the objects represented by the theory interact with our measuring apparatus. This may involve an account of the relationship between the theory at hand and other theories. If so, the ontological interpretations of all of the theories involved will be required to mesh in an appropriate way. For some theories, a full interpretation may also require an account of the role of human consciousness in the worlds described by the theory. There are surely other aspects of interpretation that play a role in philosophy of physics. Here, however, my concern will be almost exclusively with the first aspect: ontology.

It is worthwhile to run through a simple example. Let's suppose that our theory is the heat equation in one spatial dimension, , which is defined on some space of functions $X=\{u(t,x): \mathbf{R}^2 \to \mathbf{R}\}$. The standard interpretation of this theory views $u(t,\cdot)$ as representing the state of the thermal state at time t of an infinitely long (and thin!) bar. The bar is thought of as being at rest in some inertial frame, so that $u(\cdot,x)$ describes the temporal evolution of the temperature at the point labeled by x. This is a satisfactory account of the ontology of the theory—although, of course, many others could be given.

It is important to be clear from the outset about one of the conventions that is employed in discussion ontological interpretations. In the literature on the hole argument, one often finds people following Earman and Norton 1987 in querying whether two isomorphic models represent the same or different situations. At worst, I think this has led to considerable confusion—people sometimes talk as if it were somehow impossible for each of a pair of isomorphic models to represent the same situation. This is nonsense—there is no intrinsic difference between the models which could determine which model represents the given situation (see Norton 1989 for this point).[61] It is like saying that only one of two identical prints of a photograph can represent the source of the image. In our example, for instance, it is obvious that many u(t,·) can represent any given instantaneous thermal state of the bar. If the bar is very hot at a certain point, we are free to let the maximum value of u occur at any value of x. Of course, once we fix a u(t,·) which represents our given state, its erstwhile competitors can no longer be seen as representing the given state, since they must now be viewed as representing states of the bar which are qualitatively indistinguishable from our given state, but which differ from it by a translation in space.

We are tacitly fixing this sort of convention when we speak as though only one model represents a given situation; changing the convention switches the representational content of the models. All of this is worked out in considerable detail in Rynasiewicz 1994. If one is careful about the role of such tacit conventions, then Earman and Norton's way of talking is just what it should be: a harmless shorthand. In what follows, I will employ this shorthand—I will suppose that we have fixed a relation of representation between the models of our theory and the possible worlds they represent, so that when I ask whether two models which are both capable of representing the same possible world do in fact both represent it, it will be understood that I am asking a question about this fixed representational relationship, rather than a question about the representational capacity of the models. This transforms a trivial question into a substantive one. One which, as we will see, lies near the heart of the interpretative problems of gauge theories.

[61] This observation is valid only so long as we have a definite idea of what differences count as intrinsic in the context at hand. In the case of GR, we have a very definite idea—we are doing differential geometry, and so only those differences which are preserved by diffeomorphisms can underwrite differences of representational capacity.

2. How do we decide whether or not a physical theory is deterministic? One finds two sorts of answer to this question in contemporary philosophical discussions. The older one is an inheritance from logical empiricism: it makes determinism a formal property of theories considered as linguistic entities. The more recent one is unabashedly metaphysical: it takes determinism to be a matter of possible futures, and thus involves us in modal metaphysics.

Almost all recent discussions of determinism take Laplace's influential characterization as their starting point:

> We ought … to regard the present state of the universe as the effect of its anterior state and as the cause of the one which is to follow. Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. (Laplace 1951, p. 4)

Of course, this vision was by no means original with Laplace. Leibniz, for instance, maintained that "destiny consists in this, that everything is interconnected as in a chain and will as infallibly happen, before it happens, as it infallibly happened when it happened." He went on to explicate this thought in terms quite similar to those of Laplace:

> Mathematics … can elucidate such things very nicely, for everything in nature is, as it were, laid off in number, measure, and weight or force. If, for example, one sphere meets another in free space and their sizes and paths and directions before collision are known, we can then foretell and calculate how they will rebound and what course they will take after the impact. Very simple laws are followed which also apply, no matter how many spheres are taken or whether objects are taken other than spheres. From this one can see that everything proceeds mathematically—that is, infallibly—in the whole wide world, so that if someone could have a sufficient insight into the inner parts of things, and in addition had remembrance and intelligence enough to consider all the circumstances and to take them into account, he would be a prophet and see the future in the present as in a mirror.[62]

Notice that both Leibniz and Laplace are speaking about both theories and things: ideal theories which would allow you to find the future of a given the present state; and the operation of causal necessity in nature. Both are evidently quite comfortable moving back and forth between the formal and the material mode of speech—they feel no need to

---

[62] Quoted in Cassirer 1956, pp. 11-12. See Hahn 1967 and Chapter 2 of Cassirer 1956 for other pre-Laplacian formulations of Laplacian determinism. Cassirer himself had a somewhat idiosyncratic view of the importance of Laplace's formulation—he held that, despite its prevalence, it was not really taken seriously until the 1870's. See Chapter 18 of Hacking 1990 for a critique of this thesis.

distinguish between determinism as a thesis about the logical structure of ideal physical theories, and determinism as a thesis about the rigidity of the causal structure of the world.

Until the middle decades of this century, it seems, most philosophical commentators were satisfied with this sort of approach. They would quote Laplace, or perhaps paraphrase him, and then launch into a substantive discussion of the problems surrounding the notion of determinism (see, e.g., Bergson 1910, p. 181; and Peirce 1923, p. 144).[63] Under the influence of logical empiricism, however, philosophers of science began to be dissatisfied with the Laplacian characterization of determinism. It was felt that philosophy of science should be able to produce a precise analysis of determinism—one that did not depend on metaphors or on the capacities of superhuman beings. Nagel's attempt was pretty typical: "the set of laws L constitute a deterministic set of laws for S relative to K if, given the state of S at any initial time, the laws L logically determine a unique state of S for any other time" (Nagel 1961, p. 281). Here K is a set of variables, and a state S is an assignment of values to each variable in K. In Nagel's favorite example of a deterministic theory, classical point-particle mechanics, K={p,q} (to be thought of as position and momentum), S is an assignment of a value for p an q for each particle at some time t, and L is a set of differential equations depending on the forces posited.[64] Very similar definitions—varying only in degree of technical sophistication— can be found in Carnap 1958, Feigl 1953, Montague 1974, and Smart 1968.[65]

Notice that this sort of analysis of determinism concerns theories rather than things. Indeed, in an attempt to motivate his definition Nagel first formulates a definition of what it is for a "system" to be deterministic; he then cautions that "[i]t is at least potentially

---

[63] James 1896 and Russell 1953 are notable exceptions. James offered a much more metaphysical view of determinism (see Earman 1986 for discussion). Russell's analysis of the notion of determinism led him to the conclusion that the world is deterministic iff its behavior was describable by some mathematical functions. He was less than happy with this conclusion, since it seemed to make determinism trivially true.

[64] See Earman 1986 chapter III for a discussion of why this is not a good example for Nagel's purposes.

[65] Bergmann 1953 and Frank 1957 add an operationalist account of the interpretation of theories to this standard view.

misleading in suggesting that it is a *system of bodies*, rather than a *theory* about certain properties of a system of bodies, which is said to be deterministic" (Nagel 1961 p. 280). This shift of emphasis was a natural move for the logical empiricists to make: for them theories *qua* linguistic objects were the primary objects of study for philosophers of science; and it made sense to construe the modality inherent in the notion of determinism as being strictly linguistic. But it was a genuine novelty, and it had an important consequence: the logical empiricist analysis of determinism is a *formal* one. If we (for convenience) limit ourselves to theories whose laws are differential equations that govern the evolution of some set of variables in time, then the logical empiricists tell us that a theory is deterministic iff its initial value problem is well-posed.[66] Thus determinism becomes a purely mathematical property of a theory, entirely independent of its interpretation. This was a welcome conclusion for the logical empiricists, who were happy to replace the metaphors of the traditional philosophical discussion of determinism with a well-defined mathematical criterion. This made it possible, for instance, for Nagel to argue that quantum mechanics is deterministic without needing to discuss the interpretative problems of quantum theory (Nagel 1961, §10.IV).

Earman and Lewis have recently offered a very different sort of analysis of determinism—one which emphasizes worlds rather than theories, and is unabashedly metaphysical. Earman formulates a definition of a deterministic world: "Letting **W** stand for the collection of all physically possible worlds … The world $W \in \mathbf{W}$ is Laplacian deterministic just in case for any $W' \in \mathbf{W}$, if W and W' agree at any time, they agree for all times. By assumption … agreement of worlds at a time means agreement at that time on all relevant physical properties" (Earman 1986 p. 13). Lewis' definition is very similar, and

[66] The initial value problem for a set of differential equation involving kth time derivatives is said to be well-posed if, given an assignment of values to all the dependent variables and their first through (k-1)th time derivatives at a given instant of time, we have both: (i) (Existence) that there exists a solution to the equations which satisfies these initial value constraints; and (ii) (Uniqueness) that there is only one such solution. In the case of the heat equation (see §1 above), the initial value problem is well-posed if, given $u(t_0,\cdot)$, there is exactly one u(t,x) satisfying the differential equation and the initial value constraint.

has the advantage over Earman's of incorporating a precise notion of agreement (see p. 360 of Lewis 1983). This is discussed at length in the Appendix. For now it is sufficient to have in mind Earman's account of determinism as an example of the metaphysical approach.

This is clearly distinguished from the formal approach of the logical empiricists by its reliance on interpretation. For Earman and Lewis it is in the first instance a world rather than a theory which is deterministic. In order to decide whether or not a physical theory is deterministic, we must first look at the possible worlds which are correctly described by a theory; the theory is deterministic iff the worlds it describes are deterministic. Now, most theories admit of several interpretations. And in general the models of the theory will represent distinct possible worlds under distinct interpretations. This is clear in the case of quantum mechanics. Under some, but not all, interpretations, a particle possesses a certain type of property iff it is represented by a vector which is an eigenvector of a certain operator. Under some, but not all, interpretations there is a physical process of measurement corresponding to the projection postulate. The possible worlds which the models of quantum mechanics represent are very different under different interpretations. Thus it should come as no surprise that, where Nagel was able to decide that quantum mechanics is deterministic without worrying about interpretative problems, Earman is forced to withhold judgment until a satisfactory interpretation is produced (Earman 1986, Chapter XI). This situation is quite typical under the Earman-Lewis approach: questions of determinism cannot be settled in isolation from questions of interpretation.

One might hope that the approach of the logical empiricists would be equivalent to that of Earman and Lewis, despite the fact that one is formal while the other is interpretative. After all, both of them provide prima facie adequate explications of Laplace's idea. So perhaps the two approaches are equivalent, in the sense that they will yield the same judgments about which theories are deterministic. Consider the following argument for equivalence. Suppose that we have a theory which correctly describes a world. Then the old approach says that the theory i deterministic iff its initial value problem is well-posed; the new approach says that it is deterministic iff there is only one physically possible future for the world, given its present. Now, since the theory describes the world, there must be a bijection between the mathematical structure of the theory and the ontology of the world—i.e. there will be a one-to-one correspondence between initial value conditions, variables, and differential equations on the one hand, and instantaneous states, physical entities and properties, and laws of nature on the other.[67] This bijection allows us to conclude that the theory will be deterministic in the new sense iff it is deterministic in the old sense.

If this argument is correct, then there are two good reasons to prefer the old to the new. (1) Both approaches need a great deal of elaboration. But in the case of the new approach, it is highly contentious modal metaphysics which is required—we need, for instance, an explication of Earman's notion of agreement. With the old approach, the remaining work is purely technical, of the sort that Montague did so well.[68] Why make determinism more *philosophical* than it has to be? (2) At the same time, the old approach has close connections to the way that physicists think of determinism. The new approach looks like some kind of philosopher's version of what is, after all, a physicist's concern. So we have a choice between two research programmes which, we are supposing, lead to the same results. The first presents some clear and interesting technical problems. The second

---

[67] Here, of course, I assume that the language of the theory is capable of expressing uncountably many facts; see pp. 303-4 of Montague 1974. There are two ways to make this work: replace talk of languages with talk of models (Montague 1974); or work with infinitary languages (Leeds 1995).

[68] This is actually a bit misleading. Even apparently technical questions lead to metaphysical ones. For instance, it is clear that we have to put some restrictions on what counts as a "theory" if we want our definition to be interesting—otherwise we will end up, like Russell, saying that every world is deterministic, because there is *some* function which describes what happens there. A natural restriction is to require that the theory be recursively expressible in some nice language. Now consider a puzzle due to Scriven: imagine two worlds, in both of which there is nothing but a ten-sided die, whose faces light up in turn; in one world, the sequence of illuminated faces gives the decimal expansion of π, in the other world, the sequence is some non-recursively enumerable set (see p. 731 of Scriven 1957). Is the fact that there is a finite description of one but not the other really relevant to the question of determinism? Scriven thinks so; I am dubious. I suspect that one's intuitions here are dependent on one's intuitions about laws.

involves us in a difficult philosophical debate, and offers us many wrong turnings. Why would anyone who is more interested in substantive issues in the philosophy of physics than in debates about modality choose the latter course?

Notice, however, that the argument for equivalence depends crucially on the supposition that there is a bijection between the mathematical structure of the models of the theory, and the ontology of the worlds which they represent. This is a substantive assumption, and we should regard it with suspicion. It is not clear, for instance, that it will be easy to make sense of such talk in cases like quantum mechanics where the mode of representation bears little similarity to the relatively literal mode employed in classical mechanics. Furthermore, even if we can make sense of talk of bijection, we have to worry that the bijection will exist under some interpretations of quantum mechanics but not others. In this case, the Earman-Lewis approach will, presumably, return different verdicts depending on the interpretation employed, while the formal approach will, of course, not be affected by such differences. Thus, quantum mechanics may provide an example where the formal and interpretative approaches to determinism diverge. There is not, however, much hope of using this as a test case to decide between the two approaches, since we do not have in hand any satisfactory interpretations of the theory.

GR is another example where the two approaches may well fail to be equivalent. As we will see in the next Chapter, the hole argument hinges on the question of whether substantivalists are committed to viewing each point in the phase space of GR as representing a distinct physical possibility. Thus, the argument turns upon the question of the existence of a bijection between the mathematical structure of the theory and the ontology of its interpretation. So we should expect the formal and interpretative approaches to part ways over the question of the determinism of GR. It is not surprising, then, to find that much of the debate over the hole argument can be read as a clash of intuitions concerning the proper analysis of determinism. Leeds brings this out very explicitly:

> I argue that Earman and Norton's familiar "hole argument" raises questions as to whether GTR is a deterministic theory only given a certain assumption about determinism: namely, that to ask whether a

theory is deterministic is to ask about the physical situations described by the theory. I think that this is a mistake: whether a theory is deterministic is a question about what sentences can be proved within the theory. (Leeds 1995, p. 425)

If Leeds is correct, then the hole argument evaporates—the connection between substantivalism and indeterminism which it is supposed to point up can exist only if determinism is an interpretative rather than a formal property of theories. There are a number of other discussions of the hole argument which, although they do not go as far as Leeds does in explicitly endorsing the logical empiricist approach to determinism, rely nonetheless upon formal approaches to determinism: Mundy 1992, Rynasiewicz 1992, and Wilson 1993.[69] Of course, since GR is a gauge theory, and hence does not have a well-posed initial value problem in the strict sense, these authors must modify the formal approach of the logical empiricists to take gauge freedom into account. Each argues from the fact that the initial value problem of GR is well-posed upto diffeomorphism to the conclusion that the theory is deterministic. As we will see in the next Chapter, that is just to say that they take uniqueness upto gauge freedom to be sufficient for determinism, in the case of GR. Each of them has slightly different grounds for this judgment, and it is not entirely clear that they would all endorse the same criterion for determinism for other gauge theories. But it is clear that these authors advocate formal approaches to determinism, and that they are united in seeing the hole argument, with its putative connection between substantivalism and determinism, as spurious. Thus GR, like quantum mechanics, is a case

[69] There are, in addition, the hybrid approaches of Healey and Rynasiewicz. Both of these authors advocate analyses of determinism which are manifestly interpretative.

But Healey's explicit reliance on structuralist intuitions in effect introduces a formal element into his analysis (since, independent of interpretative questions, this structuralism guarantees that a theory can never be ruled indeterministic on the basis of its countenancing distinct but isomorphic possible worlds). In Healey's account, this formal element does most of the work involved in disarming the hole argument, so that any criticism of formal approaches of determinism also bears against Healey response to the hole argument (for the importance of the formal element, see especially fn. 24 of Healey 1995).

In Rynasiewicz 1994 and Rynasiewicz 1996b, on the other hand, it is difficult to see what the non-formal input into the algorithm for deciding whether or not a theory is deterministic is supposed to be. Meanwhile, it is quite clear that Rynasiewicz sees no connection between determinism and interpretative questions concerning ontology (see, e.g., Rynasiewicz 1996b, p. 1). Thus, I contend that the discussion below tells against Rynasiewicz's analysis of determinism, even though it is not a strictly formal analysis.

where there may well be is substantive disagreement between the advocates of formal and interpretative construals of determinism. But, again, there is insufficient agreement about the proper interpretation of GR for us to be able to decide the issue between the two camps by examining the details of this case.

Nonetheless, GR is a suggestive example. Gauge theories provide a promising class of examples in which to look for failures of equivalence between the formal and interpretative approaches, since the particular way in which their initial value problems fail cries out for a re-examination of the thesis that there is a bijection between mathematical structure and ontology.[70] In the next two sections, I discuss two familiar gauge theories, and argue that they not only provide examples of inequivalence between two approaches, but also demonstrate that the interpretative approach is to be preferred.

3. I am going to discuss interpretations of two familiar gauge theories, electromagnetism (example 5.7) and LB2 (example 5.6). I will present two interpretations for each of these gauge theories, and argue that in each case the theory is deterministic under one interpretation and indeterministic under the other. This, of course, suffices to show that determinism cannot be a formal property of theories. It will also show that having a well posed initial value problem is neither necessary nor sufficient condition for a theory to be deterministic.

Examples

(1) We consider the standard formulation of electromagnetism in Euclidean space, in terms of the vector potentials, A, and the electric field, E (example 5.7). Maxwell's equations take the form:

.

[70] This is not to say that the more radical varieties of failure of initial value problem considered in Earman 1986 and Wilson 1989 may not also provide interesting test case for the equivalence between the two approaches.

As discussed in Chapter 5, if we specify A and E at $t=t_0$, then we can find a solution (E(t),A(t)). But this solution is not unique. If Λ is any scalar function on spacetime, such that grad Λ(t)=0, then the pair (E(t),A(t)+grad Λ(t)) is another solution for the same initial data; and these are all the solutions for that initial data. Thus, the initial value problem for Maxwell's is not well posed in the strict sense, but is well posed upto the gauge transformation A→A+grad Λ.

According to the formal approach electromagnetism is not deterministic, since a specification of initial data does not suffice to fix a unique solution to the differential equations of the theory.

We can, however, provide interpretations of electromagnetism under which the theory is manifestly deterministic according to the Earman-Lewis approach. The traditional approach is to insist that the potentials are mere mathematical fictions, and that the gauge freedom in Maxwell's equations is a representational artifact. Typically, one stipulates that it is E and B, the electric and magnetic fields, which represent the true ontology of the theory (recall that Bcurl A). And, in the case at hand where space is simply connected, each pair (E,B) corresponds to an entire gauge orbit in the phase space constructed out of the potentials. Thus the fact that solutions to Maxwell's equations are unique upto gauge transformations means that if we fix E and B—or E and A—at a given time, then the Maxwell equations give us a unique solution for E and B for all times. Thus there is only one future compatible with a given initial state of the electric and magnetic fields, so electromagnetism is a deterministic theory under this interpretation.

There are, however, interpretations under which electromagnetism fails to single out a unique possible future for each initial state. Let's suppose that I believe in a material ether. At each point in space this ether has a velocity which, I claim, is represented by A. Correspondingly, E=– represents the acceleration of the ether.[71] Then the gauge

[71] This is only a couple of steps removed from historical reality: in Maxwellian electrodynamics the current was the time derivative of a vector field which was sometimes interpreted as representing the velocity of the ether. See Buchwald 1985, p. 24.

transformation $A(t) \rightarrow A'(t)=A+\text{grad}\ \Lambda$ is non-trivial, since the two gauge related solutions, A and A', represent very different histories for each bit of ether. Since I believe in the physical existence of the ether and its parts as material bodies, I have to admit that a possible world in which *this* bit of ether ends up *here* is distinct from a possible world in which it ends up *there*. Thus, I have to admit that electromagnetism is (Earman-Lewis) indeterministic, since specification of the initial state does not serve to single out a unique possible future.

(2) Our theory is the reduced phase space formulation of LB2. Thus we are working with the Hamiltonian system (T^*Q_0,ω,H), with Q_0 the space of relative configurations of n particles, ω the canonical symplectic structure on T^*Q_0, and H the projection of the LB2 Hamiltonian. This theory has a well posed initial value problem, and so is deterministic according to the formal approach.

We can also interpret the theory so that it is deterministic according to the interpretative approach. Under this interpretation, the possible worlds of the theory contain n point particles existing in a relationalist space. This is a viable interpretation of the theory, since the our dynamics does not require as input any spatiotemporal structure other than the instantaneous distances between bodies, and a time derivative. It is a deterministic interpretation because the bodies do not have any properties other than their instantaneous relative distances, which, given the initial data, are determined for all time by the theory.

If, however, we believe in any additional spatiotemporal structure, then the theory becomes indeterministic. Say, for example, that we believe that our spacetime has the structure of full Newtonian spacetime, so that the particles are embedded in a substantival space. Then in addition to the instantaneous relative distances, each particle has an absolute position, velocity, etc. The theory is incapable of determining these—so it is (Earman-Lewis) indeterministic.

To see how this works in detail, we first choose a more appropriate mathematical representation: the standard representation of Newtonian spacetime as $\mathbf{R}^4=\{(x^1,x^2,x^3,t)\}$. Now we will be able to talk about absolute accelerations and velocities of particles in the ordinary way. A history of the system will be represented not by a given curve in T^*Q_0 but by the imbeddings of the worldlines of the particles into $\mathbf{R}^4$ in the standard manner, so that the relative distances and velocities agree with those determined by the given curve. The problem is that there are too many such imbeddings. The theory is invariant under the group of symmetries of Leibnizean spacetime. These are maps which act on (x^1,x^2,x^3,t) by:

(here R(t) is a time-dependent rotation of space, b(t) is a time-dependent boost, c is a constant, and the summation convention for repeated indices is in effect). This means that if Φ is an embedding of the worldlines of the particles into $\mathbf{R^4}=\{(x^1,x^2,x^3,t)\}$(and hence represents a dynamically possible history), then so is $\pi\circ\Phi$, where π is a map of the form (*).

How does this get us into trouble? Fix an instantaneous state of our system, and use this as initial data to generate a unique curve in T^*Q_0. Now fix an imbedding Φ and choose π so that the c=0, R(t) is the identity for all t, and b(t)=0 for t≤0, but $b(t)=t^2(1,0,0,0)$ for t>0. As long as the initial data is posed before t=0, we have two dynamically possible histories for the same initial data: the one represented by Φ and the one represented by $\pi\circ\Phi$. These two histories do not differ as far as any relational quantities are concerned—the particles have the same relative distances and velocities in both histories. Thus there is nothing in LB2 to distinguish them. But they *do* differ by our lights, if we really believe in the full structure of Newtonian spacetime. Consider the history of a given particle under Φ and $\pi\circ\Phi$. Say the absolute velocity and acceleration of our particle under Φ are v(t) and a(t) respectively. Then under $\pi\circ\Phi$ they are (for t>0) v(t)+2t(1,0,0,0) and a(t)+(2,0,0,0) respectively. This is a real difference to us, and so we have to admit that our theory is indeterministic—the present state of the world does not serve to determine the future state.

We could equally well have taken the standard formulation of LB2 as our theory, rather than the reduced phase space formulation. Then we would be working with the gauge system (T^*Q,ω,H), where Q is the set of positions of the particles relative to a fixed

Cartesian coordinate system on $\mathbf{R}^3$. In this case, of course, the initial value problem would not be well posed, so that the theory would be indeterministic according to the formal approach. But we could still have both of our interpretations of the theory: according to the first, each gauge orbit of (T*Q,ω,H) corresponds to a single physically possible state (the set of relative positions and velocities of the particles, about which all the members of the gauge orbit agree); while according to the second, each point in the phase space corresponds to a distinct physically possible state of affairs (the positions and velocities of the particles relative to substantival space). Thus we again conclude that LB2 is deterministic according to the first interpretation, and indeterministic according to the second.

4. These examples provide illustrations of a quite common phenomenon: gauge theories admit multiple interpretations. Of course, some of these interpretations are likely to be less attractive than others. But: (i) each is, nonetheless, an interpretation which we could choose to impose on the theory; (ii) one can usually imagine circumstances under which any interpretation would become an attractive option. (For example: one might need substantival space to underwrite other physical theories, so that belief in the unity of physics would drive one to set one's interpretation of LB2 in the otherwise inappropriate setting of Newtonian spacetime.)

In the examples of the previous section, I exhibited in both cases a *literal* interpretation (in which each point in the phase space of the gauge system represented a distinct physical possibility) and a *gauge invariant* interpretation (in which each gauge orbit represented a single physically possible situation). Under a literal interpretation of a gauge theory, specification of the present state fails to fix a unique possible future, so that gauge theories are always Earman-Lewis indeterministic under such interpretations. Gauge invariant interpretations, on the other hand are always deterministic under the Earman-Lewis approach. Thus the hoped-for equivalence between the formal and interpretative approaches to determinism dissolves whenever we are dealing with a gauge theory. I believe that, in the absence of the equivalence, it is clear that it is the interpretative approach which deserves out loyalty. There is no real difference between the standard and the reduced phase space formulations of LB2, and yet the formal approach rules that one but not the other is indeterministic. It seems to me that this tells heavily against the formal approach. Furthermore, I find it untenable to maintain, e.g., that there is simply a fact of the matter about whether or not electromagnetism is deterministic, independently of what sort of ontology and ideology the theory is taken to describe.

This leaves us with the conclusion that we must reject the formal approach in favor of the interpretative approach when working with gauge theories. This will be crucial when we turn towards the project of interpreting GR. But notice that this conclusion is not of narrow scope: not only is it true, as remarked in Chapter 5, that many of the most important classical field theories are gauge theories, but it is also true that any Hamiltonian system can be transformed into a gauge system by embedding it in some larger symplectic manifold (see the discussion of gauge fixing in §5.8). This latter is not an idle technical point: in example 2 of the previous section, we began with a Hamiltonian system (the reduced phase space formulation of LB2) and gave it an indeterministic interpretation by stipulating that there were more degrees of freedom in the ontology of the interpretation than in the formalism of the theory. In the end, we saw that this was equivalent to giving a literal interpretation of the un-reduced form of the theory.

The fact is that we can always do something like this. We can embed our Hamiltonian in a larger symplectic manifold, so that it appears that our system arose as a result of gauge fixation of a constrained system (see the discussion of example 5.4). We can then give a literal interpretation of the gauge system, so that our theory is indeterministic. Alternatively, we can skip the embedding step, and simply introduce excess degrees of freedom at the level of interpretation. In this way we can give an indeterministic interpretation of any theory—even one whose dynamics is modeled by a Hamiltonian system. Thus, having a well-posed initial value problem is neither a necessary nor a sufficient condition for being deterministic.

5. In the balance of this Chapter I am going to conduct a quick survey of strategies for interpreting gauge theories. Until the final section, I will assume that we are working with a gauge theory (M,ω,H) with H≠0, so that the Hamiltonian vector fields of H are not null vector fields of ω. Thus I exclude parameterizations and other cases where there is no preferred parameterization of time. That is, I limit myself to the case where time evolution is *not* a gauge motion. Further, I restrict attention to cases where we are interested in interpretations where points of M represent instantaneous total states of possible worlds.

It is important to be clear about the relationship between these strategies for interpreting gauge theories, and the strategies canvassed in Chapter 5 for isolating a well posed initial value problem within a gauge theory. Here it is helpful to distinguish between predictability and determinism. Suppose, for instance, that I subscribe to the interpretation of electromagnetism according to which the vector potential represents the velocity of a material ether. This means that I am committed to viewing electromagnetism as an indeterministic theory—by my lights, the theory cannot determine a unique future for a given initial state. This does not mean, however, that I cannot use the theory to make definite predictions about the future. I can still exploit the fact that Hamilton's equations have a well posed initial value problem for gauge invariant functions on phase space. I can use electromagnetism to make perfectly accurate and determinate predictions concerning the evolution of physical quantities, such as the magnetic field, which are represented by gauge invariant functions. Thus I can have predictability for some quantities, without having determinism. The indeterminism of a predictable but indeterministic theory is, of course, perfectly harmless from the viewpoint of someone interested in applying the theory.

In order to mark this distinction between determinism and predictability, it is helpful to introduce a further terminological distinction, between *observables* and *beables* (the latter is borrowed from Bell, although I use it in a slightly different sense; see Essays 7 and 19 in Bell 1987). Relative to a given interpretation of our gauge theory, a function on phase space will be called an observable if it doesn't allow us to distinguish between points of phase space which represent observationally indistinguishable states of affairs. Thus if $f \in O(M,\omega)$ is an observable, and $f(x) \neq f(y)$, then x and y represent states which can be experimentally distinguished from one another. Similarly, a function on phase space will be called a beable if it doesn't allow us to distinguish between points of phase space which represent the same states of affairs. That is, if $f \in O(M,\omega)$ is a beable, and $f(x) \neq f(y)$, then x and y represent distinct physical states.

The idea is that observables correspond to properties that we can measure, while beables correspond to properties which individuate. If $x,y \in M$ represent distinct physical states, then we can find a beable f which distinguishes between x and y in the sense that $f(x) \neq f(y)$; if, on the other hand, x and y represent the same state, then $f(x)=f(y)$ for all beables f. If $x,y \in M$ represent observationally distinguishable physical states, then we can find an observable f which distinguishes between x and y in the sense that $f(x) \neq f(y)$; if, on the other hand, x and y represent states which are not observationally distinguishable, then $f(x)=f(y)$ for all observables f.

Now let $\tau x(\tau)$ and $\tau x'(\tau)$ be two dynamical trajectories with $x(0)=x'(0)$. Our interpretation is deterministic iff $f(x(\tau))=f(x'(\tau))$ for every beable f and for all $\tau \in \mathbf{R}$; it is predictable iff $f(x(\tau))=f(x'(\tau))$ for every observable f and for all $\tau \in \mathbf{R}$. Equivalently: an interpretation is deterministic iff the beables commute with the constraints, and predictable iff the observables commute with the constraints.

6. The first interpretative option is to give a literal interpretation of the gauge theory according to which every point in phase space corresponds to a distinct physical possibility, as in the literal interpretations of §3. Under such an interpretation, the theory is, of course, indeterministic. Typically, it should be easy to formulate literal interpretations. One is often, for instance, presented with a gauge theory whose phase space is a cotangent bundle T^*Q, where the configuration space Q is some space of tensors over simultaneity slices. One then stipulates that the theory represents a set of possible worlds which are populated by physical fields with the appropriate structure. This is how the interpretation of the vector potential as a physically real field is generated. In that case, it was possible to embellish the interpretation by cashing out the structure of the field in terms of the kinematic properties of an entity. With sufficient ingenuity, this is presumably often possible. But it is by no means required in order to produce a literal interpretation—I could have just as easily stipulated that the worlds represented by electromagnetism contained a physical field represented by A without worrying about providing a mechanical account of its nature, just as the traditional interpretation does with E and B.

This fixes the ontology of the theory. There is still, of course, a great deal more work to be done before the interpretative enterprise is complete. In particular, our theory lacks an account of measurement. In order to supply a complete account, we would have to show how measurement could be modeled either within the theory itself, or by considering the theory in relation to other theories. This would require a great deal of hard work. At its end, it would produce an account of which properties described by the theory are measurable. On the schematic level at which we are working, it will often be sufficient to let this end product stand in for the entire measurement theory. Thus we can think of the specification of the set of observables of the theory as a further interpretative step, beyond the specification of the ontology of the theory. It should always be understood that this specification is standing in place of a great deal of hard work—one cannot simply stipulate which quantities are observable.

Having settled on a literal interpretation for our gauge theory, the question of the size of the set of observables remains open. There are two obvious candidates for the set of observables: $GO(M,\omega)$ and $O(M,\omega)$. If we flesh out our interpretation so that the former is the set of observables, then the theory is predictable although indeterministic. If we settle on

O(M,ω), then every beable is an observable, and the theory is radically unpredictable, since the equations of motions for many beables contain arbitrary functions of time.

We could imagine either outcome in the case our literal interpretation of electromagnetism which casts the vector potential as the velocity field of a material ether.[72] If we believe that the ether is some sort of subtle, but massive, fluid which interacts mechanically with other matter, then it seems that we are going to have to admit that O(M,ω) is the set of observables—any differences in the velocity and acceleration of the fluid will be observable in principle, if not in practice. If, on the other hand, our ether is an imponderable fluid, which only interacts with ordinary matter only via electromagnetic phenomena, then we may be able to show that we can only distinguish states of the ether which give rise to distinct electromagnetic fields.

Literal interpretations which have O(M,ω) as their set of observables are quite implausible. They make the theory in question both indeterministic and unpredictable, so that they are not likely to appear very attractive so long as there are *any* alternatives available. It is hard to imagine at anyone has ever deliberately advocated such an interpretation.

Literal interpretations whose observables are gauge invariant, on the other hand, are not nearly so unattractive. Indeed, Kuchar advocates such an interpretation for GR (private communication). Since they are predictable, there is no problem in applying them to model classical systems. Put another way: their indeterminism has no straightforward empirical consequences. Thus, one need not feel uncomfortable about advocating such an interpretation, so long as one doesn't feel too squeamish about indeterminism. And, since, as we will see, it may be difficult to find adequate gauge invariant interpretations for some physical theories, a predictable literal interpretation may turn out to be the best of a bad lot. More interestingly, as we will see in Chapter 8, although predictable literal interpretations and gauge invariant interpretations of the same theory are empirically indistinguishable at

[72] Of course, any measurement theory for this interpretation is going to involve accounts of the interaction between the ether and other matter which are very different from any ever entertained by historical ether theorists.

the classical level, they may generate distinct approaches to quantization, and thus lead to empirically inequivalent quantum theories.

7. A second interpretative strategy is to fix the gauge. We have already seen how this is done at the formal level (see §5.8). There we select a constraint submanifold N of the phase space which contains exactly one member of each gauge orbit. This has the effect of eliminating the gauge freedom—since the freedom to perform gauge transformations which permute the elements of the gauge orbits is killed by our elimination of all but one of the elements of each gauge orbit. This gives us a theory with a well posed initial value problem for all $f \in O(N, \omega|_N)$, so that there are no problems with predictability.

But in order to make the theory deterministic, we must give it an interpretation under which each element of the gauge-fixed constraint surface corresponds to only one possible situation. Presumably those points of phase space which do not lie on the constraint surface defined by the gauge condition do not represent any possible situations at all. Here we must be careful to avoid re-introducing gauge freedom at the level of interpretation. Imagine for instance, that we worked with electromagnetism in the Lorentz gauge, so that the theory is cast into the form of a Hamiltonian system. It could nonetheless be indeterministic if we stipulate that the vector potential is the velocity field of a material ether, and do not recognize that this velocity field is subject to some law corresponding to the Lorentz gauge condition. If we don't recognize some such condition at the level of ontology, then two velocity fields which correspond to two vector potentials of the original non-gauge fixed formulation of electromagnetism will have equal claim to be represented by a given gauge-fixed solution. Thus, in order to insure that our theory is deterministic, we need to give the gauge conditions physical content by imposing the constraint both at the level of formalism and the level of ontology.

One must, of course, be very careful to distinguish between a gauge condition which is helpful because it renders a theory mathematically tractable, and a gauge condition which has physical content. In the case of GR, we find that if we write down the components of the EFE in some given coordinate system, $\{x^\mu\}$, then the solutions are unique only upto the addition of four arbitrary functions of time. We gain uniqueness, however, by stipulating that our given coordinate system is *harmonic*—i.e., we have that . Most authors regard this gauge choice as a mere mathematical convenience—typically the choice is explicitly compared to the physically meaningless choice of gauge in electromagnetism (see, e.g., §7.2 of Hawking and Ellis 1973 and §10.2 of Wald 1984). From this point of view, the choice of harmonic coordinates has no interpretative consequences whatsoever. Fock, however, believed that in the asymptotically flat case harmonic coordinates reflected physical reality more accurately then other coordinates. Thus, he wanted to define what it meant for a line to be straight by using linear equations in terms of the privileged coordinates rather than in terms of the affine connection on spacetime; and he wanted to use harmonic coordinates to draw an observer-independent distinction between gravitational and inertial acceleration (see pp. 224 and 232 respectively of Fock 1964). The skepticism with which Fock's claims are generally met is a good measure of the difficulty of employing gauge fixation as an interpretative strategy.

8. The final option which we will consider is to select a gauge invariant interpretation. Here each gauge orbit of the gauge system corresponds to a single physical possibility, so that the theory is deterministic. If we are working with the full gauge theory, then the beables will be the gauge invariant functions. We can also, of course, work with the reduced phase space, and give it a literal interpretation.

Gauge invariant interpretations are very attractive: they render the theory predictable and deterministic, without requiring the imposition of arbitrary-seeming constraints. Thus it might appear that every gauge theory should be given a gauge invariant interpretation. Perhaps that is the case. Unfortunately, it is not always easy to *find* a gauge invariant interpretation. The problem is that many gauge theories are based on configuration spaces which are naturally regarded as spaces of possible configurations of some system (a field or a set of particles) in Euclidean space. This makes it easy to offer literal or gauge fixed interpretations of the theory—just posit the existence of fields or particles which have the same structures as those of the mathematical representation. But a gauge invariant interpretation is a literal interpretation of a reduced phase space. And it is not always easy, or even possible, to view the reduced phase space as representing the instantaneous dynamical states of a field or a set of particles in space.

I illustrate this problem using two familiar examples. It is important to note that these are *not* examples where technical difficulties prevent us from constructing a reduced phase space formulation of the theory, or where the reduced phase space fails to be the cotangent bundle of a nice configuration space (see §5.8.iii). Rather, they are examples where the configuration space upon which the reduced phase space is based is not the space of configurations of a set of particles or fields relative to physical space. Thus it requires some work to find a literal interpretation of the reduced phase space formulation.

Examples.

(3) Consider non-relativistic theories of gravitating point particles. The standard Newtonian theory can be formulated as a Hamiltonian system (T^*Q,ω,H), with Q the space of possible configurations of the n particles relative to some system of Cartesian coordinates on $\mathbf{R}^3$. We can interpret this theory using an ontology consisting of n particles existing in substantival space, so that the coordinate system upon which Q is based is a fixed set of coordinates on substantival space. Then each point in T^*Q corresponds to a distinct possible dynamically state of the physical system, and the theory is deterministic.

The theory LB2 (example 5.6) is a gauge theory if we take the space Q of the Newtonian theory as our configuration space. If however, we begin with the space Q_0 of relative configurations (the space of possible relative distances between the particles), then we can construct the reduced phase space of the gauge system (the construction results in a

Hamiltonian system). We can interpret either the gauge system formulation or the reduced phase space formulation as being a theory of the gravitational interaction of n particles existing in a relational space.

But if we consider the intermediate theory, LB1 (example 5.5), then the situation is less straightforward. With LB2, we built the configurations space of the reduced theory, Q_0, by identifying points in Q which differ by the action of an orientation-preserving Euclidean isometry. In LB1 the configuration space of the reduced theory is built by identifying points of Q which differ by translations in Euclidean space (so that points which differ by a rotation are still considered to be distinct). This means that LB1, which admits a standard of absolute rotation but not of absolute linear acceleration, stands between the Newtonian theory with its full complement of notions of absolute motion, and LB2 which recognizes no notion of absolute motion.[73] As a consequence, the three dimensional substantivalist interpretation which perfectly fit the Newtonian theory and the three dimensional relationalist interpretation which perfectly fit LB2 are both inappropriate as interpretations of LB1. On the one hand, a three dimensional substantivalist interpretation of LB1 would be indeterministic, just as in the case of LB2 (see example 2). The same objection applies to any four dimensional substantivalist interpretation which implies that the Leibniz-Clarke counterfactuals are true. On the other, straightforward three dimensional relationalism is not a viable interpretative option for LB1, any more than for the Newtonian theory—since both theories recognize varieties of absolute motion.

At this point, we can revert to the strategies discussed in Chapters 2 and 3. We can perform Sklar's maneuver, by supplementing the three dimensional relationalist ontology and ideology with a primitive property of absolute rotation. This gives us a viable three dimensional interpretation—it is perhaps, somewhat uncomfortable, but is likely to appear more attractive than the indeterministic three dimensional substantivalist interpretation.

---

[73] Notice the similarities with the Newton-Cartan theory. I believe that the points made below concerning the interpretative options for LB1 also go through for the Newton-Cartan theory.

We can also give the theory a four dimensional relationalist interpretation: we let spacetime be the set of possible spatiotemporal relations between worldlines, and build exactly enough structure into this set to underwrite notions of spatial distance between simultaneous slices of worldlines, and a notion of rotation for each individual worldline. This leads to a viable interpretation of the theory.

So we have the following interpretative options for gauge invariant interpretations of LB1: (i) three dimensional relationalism with Sklar's maneuver; (ii) four dimensional relationalism with spacetime conceived of as being the appropriate geometry of possible spatiotemporal relations between worldlines; (iii) some form of Lockean four dimensional substantivalism, according to which the Leibniz-Clarke counterfactuals are meaningless.

(4) We return once more to electromagnetism. If we continue to work with ordinary three dimensional Euclidean space, then we have the following situation. We begin with a configuration space Q={A:$\mathbf{R}^3 \rightarrow \mathbf{R}^3$} of vector potentials on Euclidean space. We then build the cotangent bundle T*Q, and look at the constraint surface N on which the momentum canonically conjugate to A (i.e., the electric field, E) is required to have zero divergence. This gaves us a gauge theory. It turns out that the reduced phase space of the theory is just the set of pairs (E,B) of divergence free magnetic and electric fields on Euclidean space. There is a bijection between the reduced phase space and the gauge orbits of the original gauge theory. So specifying a point (E,B) in the reduced phase space serves to fix a class of gauge related vector potentials. If, for instance, we stipulate that B=E=0, then we single out the gauge orbit {(A,E): where A_0=0} (i.e., every vector potential that gives rise to B=0 can be gauge transformed to the zero potential). We can give a literal interpretation of this reduced phase space formulation, according to which there are physical electric and magnetic fields. These are defined on each simultaneity slice of spacetime, and their structure is accurately described by Maxwell's equations. This interpretation is, of course, deterministic.

Thus, it appears that electromagnetism provides a good example of a gauge theory which admits an obvious and perfectly satisfactory gauge invariant interpretation. This appearance dissolves, however, if we admit topologically non-trivial spacetimes (specifically: spacetimes which are multiply connected). Consider, for instance, a spacetime whose spatial slices are of the form $M=\mathbf{R}^3/\{(0,0,z): z\in\mathbf{R}\}$ (i.e., delete the z-axis from $\mathbf{R}^3$). Then the configuration space of the gauge theory is $Q'=\{A: M\rightarrow\mathbf{R}^3\}$, the set of vector potentials on M. We can go on to construct a gauge theory of electromagnetism set in T*Q', as in example 5.7. But now something curious happens: the reduced phase space is not just the set of divergence free electric and magnetic fields on M. Specification of a divergence free pair (E,B) does not suffice to single out a gauge orbit in the phase space of the gauge system. In particular, there are a number of a gauge orbits which correspond to the condition B=E=0. That is, we can have curl A=curl A'=0 without there being a Λ such that A'=A+grad Λ. There are vector potentials A which cannot be gauge transformed to be zero, but which correspond to null magnetic fields.

The physical manifestation of this phenomenon is the Aharonov-Bohm effect (Aharonov and Bohm 1959; see also Aharonov 1986 for an account and references). We suppose that we have a solenoid which is long compared to the wavelength of the particles that we are dealing with, so that we can treat it in our model as being an infinitely long cylinder. We suppose that the solenoid is well insulated, so that the exterior electromagnetic field is zero no matter whether the machine is on or off.

We are also interested in the behavior of the vector potential in the exterior region. We cannot, of course, directly measure the value of the potential (since it is not a gauge invariant quantity, and so is not observable under any gauge invariant interpretation of electromagnetism). We can however, measure the exponential of the integral of the potential over a closed curve: this quantity, the *holonomy*, is gauge invariant, and corresponds to the rotation in phase which a particle experiences as it is transported along the curve.

There is no phase rotation for any closed curve when the solenoid is off—all of the holonomies are zero. When the solenoid is on, however, we find phase rotation if we transport the particle around the solenoid. That is, any loop enclosing the solenoid has non-zero holonomy (homotopic curves have the same holonomy when the electromagnetic field is zero, so all that really matters is how many times the curve loops around the solenoid).

We model this phenomenon by using $M=\mathbf{R}^3/\{(0,0,z): z\in\mathbf{R}\}$ to represent the region external to the solenoid, and using the potential A=0 when the solenoid is off and some other potential A' when the solenoid is on. We can always find an A' which gives rise to the experimentally observed holonomies and which corresponds to E=B=0. Clearly, A and A' are not related by a gauge transformation of the form $A\rightarrow A+\text{grad }\Lambda$.

In order to specify a physical state—or, equivalently, a gauge orbit in T*Q'—we must specify the holonomy around the solenoid as well as the exterior electric and magnetic fields. Thus we are faced with an unpleasant situation. It is easy to offer a literal interpretation of the full gauge theory, by reading A as representing a physical field—but of course such an interpretation is indeterministic. A literal interpretation of the reduced phase space, on the other hand, would be deterministic. But it is not straightforward to find such an interpretation. It is, as in the Minkowski case, easy enough to read E and B as physical fields. But each point in the reduced phase space contains, in addition to information about the values of E and B, information about holonomies. But it is hard to see how this additional information is to be understood physically, since it corresponds to global rather than local information. Thus it cannot correspond to an ordinary field on spacetime.

There are a number of obvious interpretative strategies for dealing with this problem, none of which are entirely satisfactory:

(i) Refuse to take the problem seriously. After all: no solenoid in a laboratory is infinitely long. So, there must be more physically realistic models of the Aharonov-Bohm effect which do not rely upon the trick of pretending that space has a non-trivial topology. In such models, presumably, we can explain the phase rotation of the particles by looking at the electromagnetic field in the interior of the solenoid. Thus we can reduce all known electromagnetic effects to features of the electric and magnetic fields. Of course, we have to admit that these effects are non-local, since the phase rotation of the particle cannot be explained by looking at the values of the fields at points along its trajectory. But this may appear more attractive than the alternative: admitting that the vector potentials have some physical content without directly representing real fields. Unfortunately, this response is only viable so long we can reasonably insist that spacetime itself has a trivial topology…

(ii) Geometrize. Electromagnetism can be formulated as the theory of a connection on a U(1) bundle over space time (i.e., the theory of the affine structure of the manifold formed by attaching to each point of spacetime a copy of the group of unitary transformations on a one dimensional complex vector space). From this point of view, the electromagnetic field tensor represents the curvature of the geometry. The Aharonov-Bohm effect then amounts to the fact that one in general expects that parallel transporting a vector around a closed curve will result in some non-trivial transformation, even when the geometry is flat (this can occur, for instance, in the case of a flat Riemannian geometry on a torus). Cao 1988 considers taking this formalism seriously at the level of interpretation, by hypostasizing the bundle—so that the fundamental ontology of our theory would consist of a physically real five dimensional geometry, and spacetime would be demoted to merely derivative status. This is a powerful interpretative strategy, since it can also be applied to the classical theories of strong and weak interactions (which, like electromagnetism, are Yang-Mills theories). But this strength is somewhat disconcerting, since the incorporation of each force leads to an increase in the number of dimensions of the fundamental geometrical structure of the world.

Of course, this geometrical approach is likely to strike some as being ontologically extravagant. Surely *any* physical property can be modeled by adding extra dimensions (corresponding to color, pitch, or whatever) to each spacetime point. Why take gauge potentials more seriously than these other properties?[74]

Note also that there are a number of interpretative stances which one can take towards the higher dimensional geometries of gauge theories. In particular, the substantival-relational dialectic can be played out at this level. It seems to me that fiber bundle

[74] See §2.2 of Kennedy 1993 a version of this objection.

substantivalism is exactly as vulnerable to the hole argument as substantivalism about the spacetime of GR is.[75]

(iii) Dispositionalize. Kennedy 1993 advocates reading the vector potential of electromagnetism as encoding information about the gauge independent, but path dependent, potential momentum of a particle in an electromagnetic field. Thus, the vector potential tells us how much momentum *this* particle *here* would gain if we moved it over *there*, *thus*. In order for this to count as an attempt at interpretation, in the ontological sense in which I am interested, we would need to supplement this with an account of how this path dependent property could be understood as an assignment of complicated dispositional properties to spacetime points. If this could be accomplished, then we could view a gauge orbit of potentials as corresponding to a configuration of a physical field on space.

(iv) Loops. We can recast electromagnetism as a Hamiltonian system if we begin with a configurations space which is a space of scalar functions on the space of closed curves in physical space (see Loll 1994; the scalar is closely related to the holonomies considered above). We can either: (a) as in (iii), attempt to interpret this as an assignment of dispositional properties to points of space; or (b) view electromagnetism as a theory which assigns properties to closed curves in space. This latter interpretation leads to a sort of non-local generalization of the field concept, which is a radical departure.[76] Again, however, this may be useful in interpreting other physical theories, including gravity (see, e.g., Ashtekar

[75] Textbook presentations of gauge theories tend to motivate gauge invariance by appealing to passive gauge transformations (see, e.g., §7.11 of Nash and Sen 1983 and §4.2 of Göckeler and Schücker 1987). Consideration of active gauge transformations raises the specter of indeterminism for electromagnetism, just as the consideration of active diffeomorphisms does in the case of GR.

[76] I have the intuition, which I have been unable to find a cogent justification for, that this sort of non-locality is more compatible with relationalism than substantivalism.

1995 or Baez 1994 for a discussion of the role of loops in contemporary research on quantum gravity).

(v) Radicalism. Recognizing the difficulty of the task of finding a gauge invariant interpretation of electromagnetism as a theory of physical fields on space, Brown 1994 offers an interpretation according to which the vector potential is an *abstract* object.[77]

9. So far we have restricted attention to theories for which time evolution is not a gauge motion, and concentrated on interpretations according to which a point in phase space corresponds to an instantaneous physical state of a system. We have identified three interpretative strategies:

(i) Opt for a literal interpretation. Such an interpretation is necessarily indeterministic. But this indeterminism need not have any operational content at the classical level: if the observables of the interpretation are represented by gauge invariant functions then the gauge theory is perfectly predictable. Indeterminism is, nonetheless, widely supposed to be an unattractive feature for an interpretation of a classical theory. Such interpretations are often much easier to come by than gauge invariant interpretations.

(ii) Fix the gauge. Here, aside from the technical difficulties discussed in Chapter 5, we must be prepared to face the objection that gauge fixing is an arbitrary procedure. The validity of this objection will, of course vary from case to case. This option, like the previous one, is often easy to carry out once the technical problems have been overcome.

(iii) Look for a gauge invariant interpretation. This is surely the most aesthetically pleasing solution—it gives us a deterministic interpretation without introducing any element of arbitrariness. Unfortunately, however, it is not always easy to find an appropriate ontology.

[77] In Brown's own words: "The disadvantage is obvious. I have explained the obscure and paradoxical ontology of the vector potential by appeal to an even more obscure and paradoxical ontology of quantum states, abstract entities and laws of nature—hardly a clarification of the issue. Conceded—but I am unrepentant." Brown 1994, p. 159.

10. In this final section, I turn to time reparameterization invariant theories, in which H=0 so that time evolution is a gauge motion. Each of the three options surveyed above remains a viable approach in this context. But we must take some care in applying them in this sort of case.

(i) Literal interpretation. This approach has the advantage that if $x,y \in M$ lie on the same dynamical trajectory/gauge orbit, then they are treated as representing distinct states, so that time evolution leads to genuine change. In particular, we can have an observable f with $f(x) \neq f(y)$, so that time evolution is observable. This comes at a price, however. As always, when we give a literal interpretation of a gauge theory, we have to live with indeterminism. But here the indeterminism is very peculiar—it reflects the possibility that time could flow at different rates. Indeed, suppose that our gauge system is a parameterization of an ordinary Hamiltonian system. Then if $\tau x(\tau)$ and $\tau x'(\tau)$ are two dynamical trajectories of the gauge system with $x(0)=x'(0)$, we know that both trajectories correspond to the same one parameter family of points in the phase space of the unparameterized system. Thus, we can think of our two dynamical trajectories as representing the physical system as passing through the same series of physical states. Whence, then, the indeterminism? From the fact that the two trajectories portray the system as passing through that system at different 'speeds.' A given initial state determines the order of the states that will follow it, but not the rate at which the flow of time will carry us through those states.

(ii) Gauge fixing/ deparameterization. Here we give a literal interpretation of the gauge fixed/deparameterized system according to which each point in phase space represents a distinct instantaneous physical state. We now have determinism *and* the possibility of observable time evolution. However, as always, we have to worry about possibility that the choice of gauge is arbitrary. At the very least, the time parameter should be singled out using only features which are intrinsic to the gauge system, rather than being imposed from outside (see Smolin 1991 for this point of view). As we will see in our discussion of quantum gravity, however, this is by no means a sufficient condition for privileging a choice of time at the interpretative level.

(iii) Gauge invariance. Here each gauge orbit must be taken to represent a complete history of the physical system. This is an elegant approach. One of the disadvantages is that each observable must be a constant of motion, since it is constant on each gauge orbit. This makes the gauge theory appear to be timeless: there is no evolution (since each point along a given dynamically trajectory represents the same state); and the observables allow us to distinguish between distinct dynamical trajectories, but not between distinct points along a given trajectory.

One way to understand this framework is to notice that the observables, although constants of motion, need not correspond to questions like, “what is the total angular momentum of the system?”, which mention only conserved quantities. They can correspond to questions like: “what was the rocket’s mass when it collided with the asteroid?”, which mention non-conserved quantities, as well as a time at which they are to be measured. Although the mass-of-the-rocket changes over time, the mass-of-the-rocket-when-it-collided-with-the-asteroid never changes. Thus the latter quantity is a constant of motion.

Rovelli has made the following proposal for finding dynamics and change within the timeless framework of gauge invariant interpretations of time reparameterization invariant theories.[78] In cases where our time reparameterization invariant theory is a parameterization of an ordinary Hamiltonian system, each observable of the Hamiltonian system corresponds to a one parameter family of observables of its parameterization. Thus, if the mass, m, of the rocket is an observable of the Hamiltonian system, then there is a one parameter family, $\{m(t)\}_{t\in \mathbf{R}}$ of constants of motion on the parameterization which correspond to the mass of the rocket at different times in the Hamiltonian system: m(0) is the observable “the mass of the rocket when it collided with the asteroid,” m(1) is “the mass of the rocket when it

[78] See Rovelli 1991d and Rovelli 1991a. See Hajicek 1991 and Kuchar 1993a for critical discussion; Rovelli 1991b is a response to Hajicek.

docked at the space station", etc. Now we can see change and evolution encoded in this division of the set of constants of motion into *evolving constants of motion*. The hope is that this program can be carried out even when our time reparameterization invariant theory does not arise naturally as a parameterization of a Hamiltonian system. That is, one would like to find a way of talking about observables as falling into families of evolving constants that does not rely upon the theory possessing a 'natural' choice of time variable. We will occasion to return to this suggestion in the course of our discussion of observables in GR.

CHAPTER 7
General Relativity as a Gauge Theory: The Hole Argument

1. Ideally, we would like to have a Hamiltonian formulation of GR in which the points in phase space represented instantaneous states of the gravitational field. Unfortunately, the general covariance of the theory renders this goal unobtainable. Part of the content of general covariance is that there is no preferred time coordinate in GR. Thus the theory is time reparameterization invariant. If we look for a phase space formulation of GR in which the points of phase space represent instantaneous states of the system, we end up with a gauge theory with a Hamiltonian equal to zero—so that time evolution is a gauge motion. This gauge theory has a reduced phase space which is a genuine Hamiltonian system. But, as in §5.9, each point in this reduced phase space corresponds to a set of gauge related dynamical trajectories; that is, each point in the reduced phase space corresponds to a history of the system rather than to an instantaneous state. Thus it is impossible to give a Hamiltonian formulation of GR in which the points of phase space correspond to instantaneous states of the system without breaking the general covariance of the theory.

The discussion of the previous Chapter leaves us with two conclusions: (i) every gauge theory is indeterministic under its literal interpretations; (ii) it is not always easy to find attractive gauge invariant interpretations which make the theory deterministic. The goal of this Chapter is to explore the ramifications of these conclusions for the special case of GR. This serves two purposes. First it allows us to bring the apparatus of the previous two Chapters to bear on the substantival-relational debate. This, in turn, allows us to complete our survey of the bearing of intratheoretic considerations on the interpretative problems of GR. We do this by facing a question which has been bracketed since Chapter 4—the relevance of general covariance to the interpretative project. Second, recasting the interpretative problems of GR in terms of the apparatus developed in the previous two Chapters allows me to go on in the next Chapter to establish a point of contact between the hole argument literature and the conceptual problems of quantum gravity. It is here, at last, that intertheoretic considerations bear upon the substantival-relational debate.

In the next section, I present GR as a gauge theory. I then turn to (intratheoretic) interpretative questions. §3 presents an overview of my conclusions, while §§4 and 5 provide detailed catalog of the possible substantivalist and relationalist interpretations of GR. Finally, in §6 I turn to the questions of the relative merits of the available interpretative options. Here my stance is that each of the interpretations is viable from the standpoint of the classical theory. I do, however, argue that some interpretations are preferable to others from this standpoint. I briefly discuss the relevance of the interpretative interplay between classical and quantum gravity to the dialectic which has played itself out in recent philosophical discussions of the hole argument.

2. In this Chapter we will limit discussion to globally hyperbolic vacuum solutions of GR with compact Cauchy surfaces. As a rule of thumb, one can think of the content of these restrictions as follows. The restriction to globally hyperbolic spacetimes is substantive—much of what follows is simply false, or poorly understood in the non-globally hyperbolic case. The restriction to vacuum solutions, on the other hand, is purely for convenience's sake. Almost all of what will be said is true when matter fields are taken into consideration—although the formalism involved is often more unwieldy if matter is included. The restriction to spatially compact spacetimes lies somewhere between these two extremes. There are some interesting and important differences between the compact and the asymptotically flat cases (see Ashtekar, Bombelli, and Reula 1991 for a discussion of the latter). But, for the most part, taking these differences into accounts would involve adding many qualifications to our technical treatment, without substantially altering the interpretative theses that I defend.

I also note that for the duration of this Chapter I restrict myself to the full theory of GR, without contemplating any gauge fixation, or other techniques to break the general covariance of the formalism. Furthermore, I will make heavy use of a particular form of the constraints which determine the phase space of GR. This is unavoidable—the constraints used are the only ones available. Much of the discussion will hang one whether the beables observables of one or another interpretation of GR commute with these constraints. Thus it is clear that considering alternative forms of the constraints would be helpful in surveying t the other hand, as we will see, the standard constraints are already quite well suited for our purposes. Thus, the concentration on a single form of the constraints, although less than ideal, is nonetheless not a major drawback of my approach.

We begin our search for a gauge theoretic formulation of GR by considering how to represent an instantaneous state of a general relativistic world, since we will want to work with the set of such representations as our phase space. To this end, we fix for of the remainder of this section, a compact three manifold, Σ. Now consider a globally hyperbolic vacuum solution of the EFE, (M,g), whose Cauchy surfaces are diffeomorphic to Σ. We now embed Σ in (M,g) via a diffeomorphism $\phi: \Sigma \rightarrow M$ such that $S=\phi(\Sigma)$ is a Cauchy surface of (M,g), and study the geometry which g induces on S. This geometry is characterized by two symmetric tensors on S, q_{ab} and K_{ab}. q is the Riemannian metric on S, called the *first fundamental form*, which results from restricting to T_xS (for $x \in S$) the inner product which g induces on T_xM. K is the *second fundamental form*, or *extrinsic curvature,* which encodes information about how S is embedded in (M,g). *Very* roughly, the extrinsic curvature of S is the time derivative of q (see equation 10.2.13 of Wald 1984). We use ϕ to pull these tensors back to Σ, and henceforth regard them as being defined on Σ rather than on S.

This tells us what sort of geometric structure Σ inherits when viewed as a submanifold of (M,g). Now suppose that we imagine Σ to come equipped with symmetric tensors q and K, with q a Riemannian metric on Σ. It is natural to wonder under what circumstances we can view (Σ,q,K) as being the geometry of a Cauchy surface of some model (M,g). The answer is that Σ may be embedded in a (M,g) in such a way that q and K arise as the first and second fundamental forms of Σ iff the following two equations hold:

(Gauss)

(Codazzi)

where the metric q on Σ is used to define the scalar curvature, R, and the covariant derivative, ∇.[79]

All of this suggests that we should regard a pair (q,K) as representing the dynamical state of gravitational field at a given time iff it satisfies the Gauss and Codazzi constraints. The metric q describes the geometry of a Cauchy surface; the symmetric tensor K describes the embedding of the slice in the ambient spacetime, and corresponds roughly to the time derivative of q. Thus, we can regard q as the 'position' of the gravitational field, and K as its 'momentum.'

The natural starting point for writing down GR as a constrained Hamiltonian system is Riem Σ, the space of Riemannian metrics on Σ. We regard this as the configuration space, Q, of our theory of gravity. In order to construct the phase space, we first construct T*Q, and endow it with the symplectic structure, ω, which is induced by stipulating that

is the momentum canonically conjugate to q. We then take as our phase space the constraint surface $N \subset T^*Q$ given by the following two sets of first class constraints:

(scalar)

(vector).

Each of these equations determines an infinite dimensional family of constraints, since each of them must hold at every point of Σ. Notice that the scalar and vector constraints are just

[79] Notice that these conditions make reference only to q and K—they do not mention g.

the Gauss and Codazzi constraints, rewritten in terms of p rather than K. Let $\sigma = \omega|_N$, and let H0. Then GR is the gauge theory (N,σ,H).[80]

At each point x∈N, dim ker =∞. Thus the gauge orbits of (N,σ,H) are infinite dimensional. These orbits have the following structure. Fix x=(p,q) and x'=(p',q') in N. Then x and x' lie in the same gauge orbit iff there is a solution to the EFE, (M,g), and embeddings, ϕ,ϕ': $\Sigma \to M$, such that: (i) $\phi(\Sigma)$ and $\phi'(\Sigma)$ are Cauchy surfaces of (M,g); (ii) q and q' are the first fundamental forms of $\phi(\Sigma)$ and $\phi'(\Sigma)$; (iii) p and p' are the second fundamental forms of $\phi(\Sigma)$ and $\phi'(\Sigma)$. That is, two points are gauge related iff they can be viewed as describing spatial geometries of the same model of GR. Thus each gauge orbit can be viewed as being the space imbeddings of Σ as a Cauchy surface of some model (M,g) (it could, of course, equally well be viewed as being the space of such imbeddings for any other model isometric to (M,g)).

Given a model $\mathcal{M}$=(M,g), we can find a dynamical trajectory of (N,σ,H) corresponding to $\mathcal{M}$ as follows. We first choose a foliation of $\mathcal{M}$ by Cauchy surfaces (which are, of course, all diffeomorphic to Σ). We then choose a time function τ: $M \to \mathbf{R}$, which is compatible with the foliation in the sense that the level surfaces of τ are the Cauchy surfaces of the foliation. Finally, we choose a diffeomorphism Φ: $M \to \Sigma \times \mathbf{R}$ such that each Cauchy surface of the foliation, S, is mapped onto a set of the form $\Sigma \times \{\tau\}$, where τ is the value of the time function on S. We call such a diffeomorphism an *identification map*, since it gives us a way of identifying the leaves of the foliation with Σ. We use Φ to drag along g, so that the $\mathcal{M}' = (\Sigma \times \mathbf{R}, \Phi^* g)$ is isometric to $\mathcal{M}$; the surfaces $\Sigma \times \{t\}$ in $\mathcal{M}'$ are Cauchy surfaces isometric to the Cauchy surfaces of our preferred foliation of $\mathcal{M}$. Now let $q_{ab}(t)$ and $p_{ab}(t)$ be the first and second fundamental forms of the Cauchy surface $\Sigma \times \{t\}$ in $\mathcal{M}'$. As t varies, $q_{ab}(t)$ and $p_{ab}(t)$ sweep out a curve in N. This curve is a dynamical trajectory of (N,σ,H).

[80] I will not attempt to justify this assertion here. The standard justification begins with the L d on the Einstein-Hilbert action, and proceeds via the Legendre transform and the Dirac-Bergmann algorithm. For the details of this approach see, e.g., Beig 1994 or Appendix E of Wald 1984.

Choosing a different foliation, time function or identification map will gives us a new dynamical trajectory, which will be related to the first by a gauge transformation.

This shows that, despite our emphasis on spatial geometry, our approach need not violate the general covariance of GR. If we regard gauge transformations as being physically meaningless, then our theory is indifferent to changes of foliation, time function, and identification map.

Each constraint generates a Hamiltonian vector field on N; each such Hamiltonian vector field is a field of null vectors of σ. Suppose that we look at two points x=(p,q) and x'=(p',q') which lie in the same gauge orbit, and which can be joined by an integral curve of a vector field generated by the vector constraint. Then we find that there is a diffeomorphism d: $\Sigma \rightarrow \Sigma$ such that d*q=q' and d*p=p'.[81] That is, we can regard x and x' as agreeing on the geometrical structure of Σ, and disagreeing only as to how the underlying geometrical properties are shared out over the points of Σ—x and x' may represent, for example, a geometry on Σ which has a single point of maximum scalar curvature, but according to x this point is $z \in \Sigma$, whereas according to x' it is $z' \in \Sigma$. Thus we can view the gauge transformations generated by the vector constraint as shuffling the geometrical roles played by the points of Σ. Unfortunately, the gauge transformations generated by the scalar constraint are considerably more complex.[82] *Very* roughly, they can be thought of as corresponding to time evolution—two points differ by a gauge transformation generated by the scalar constraint if they can be seen as representing distinct Cauchy surfaces in a given

[81] Notice than an integral curve of a vector field generated by a constraint is a dynamical trajectory, since H=0. But such a dynamical trajectory need not correspond to a model, (M,g), of the ordinary formulation of the theory. In the case at hand, a dynamical trajectory generated by the vector constraint gives us a one parameter family of states (p(t),q(t)). But in general it is impossible to stack the geometries represented by these states to form a model of GR: any pair of states on the trajectory are related by a diffeomorphism of Σ; that is, they represent the *same* spatial geometry; but in general it will be impossible to stack a series of copies of a given geometry to form a model of GR (not every spatial geometry is the geometry of a Cauchy surface of a stationary spacetime).

[82] This is because the action of the gauge transformations generated by the vector constraint can be viewed as arising from the action of a Lie group on the phase space of GR, while the action of the gauge transformations generated by the scalar constraint cannot be so viewed (see §4.3.5 of Isham 1991).

model. In a generic spacetime, distinct Cauchy surfaces can be expected to have very different geometries, so that points in N which are related by a gauge transformation generated by the scalar constraint will not in general represent the same geometry. In general, of course, a given gauge transformation will be generated by *both* sorts of constraint.

Now suppose that we have two dynamical trajectories which correspond to the same model (M,g). Suppose, further, that the trajectories differ by a gauge transformation generated by the vector constraint. Then, in terms of the construction above which establishes a correspondence between models and dynamical trajectories: we can use the same foliation by Cauchy surfaces and the same time function τ to generate both dynamical trajectories; the difference between the trajectories can be attributed solely to the freedom available in the choice of an identification map. If, on the other hand, the trajectories differ by a gauge transformation generated by the scalar constraint, then the difference can be traced to the freedom in the choice of foliation and time function on (M,g).

It is, as always, possible to convert this gauge theory into a true Hamiltonian system by factoring out the action of the gauge transformations to construct the reduced phase space (see §5.8.iii). In the case of GR, it is illuminating to proceed in two steps: we first partially reduce the phase space by factoring by the action of the gauge transformation generated by the vector constraint; we then complete the reduction by removing the gauge freedom associated with the scalar constraint.[83]

At the first stage, we identify any two points in N which are related by a gauge transformation generated by the vector constraint. The partially reduced phase space which results can be constructed as follows. We return to the beginning of our construction of (N,σ,H), and replace the configuration space $Q=\text{Riem}\ \Sigma$ of metrics on Σ by

$$Q_0=\text{Riem}\ \Sigma/\ \text{Diff}\ \Sigma,$$

[83] The different roles played by the two constraints in this construction can be traced back to the fact noted in the previous footnote.

the set of equivalence classes of diffeomorphically related metrics on Σ, called *superspace*. That is, superspace is the space that results from identifying two points $q,q' \in Q$ iff there is a diffeomorphism $d: \Sigma \to \Sigma$ such that $q'=d^*q$. We now look at T^*Q_0 and impose the scalar constraint and a Hamiltonian H=0 to get a gauge theory (N',σ',H). This gauge theory is the partially reduced phase space formulation of GR. By identifying diffeomorphically related three metrics from the start, we have eliminated the need for the vector constraint.

The gauge orbits remain infinite dimensional even after this partial reduction has been carried out. So we cannot apply the strategy of deparameterization discussed in §5.9 (since deparameterization requires that the gauge orbits be one dimensional). But we can still go on to eliminate the gauge freedom associated with the scalar constraint. If we identify points in N' which are related by a gauge transformation generated by the scalar constraint, then we end up with a Hamiltonian system (M,ω,H), where points in the phase space correspond to equivalence classes of diffeomorphically related models of GR. This should be no surprise: since the theory is time reparameterization invariant, the reduced phase space of GR is the space of solutions—just as in §5.9.

3. I now turn to questions of interpretation. In broadest outline, I will be running a version of the hole argument: I will attempt to show that substantivalist interpretations of GR are literal interpretations, and hence are indeterministic. I do not, however, regard this as being a reason to reject substantivalism out of hand. In the next Chapter, I will argue that this indeterminism can be a boon for substantivalism, since it means that substantivalists can consider certain options for quantizing gravity which are not available for relationalists. I conclude that the substantival-relational debate is closely related to strategies for resolving some of the central conceptual problems of quantum gravity.

The fine-grained picture is somewhat more subtle. Because it can be hard to keep track of the details, I begin this discussion by providing an overview of the interpretative options for GR. I first translate some of the terminology of previous Chapters into the General Relativistic context.

In Chapters 2 and 3, the Leibniz-Clarke counterfactuals took the form: “if all spatial (spatiotemporal) relations between bodies (material events) were as they are, but everything were shifted three meters to the left, it would be a different world.” At this point, it is helpful to consider a second formulation: “there is a possible world which is qualitatively identical to this one, but in which the points of space (spacetime) have different properties.” The idea behind the new formulation is that the material bodies of these two worlds stand in the same relations (so that the worlds are qualitatively identical), while bodies occupy different parts of space (spacetime) in the two worlds (so that the points of space/spacetime have different properties). The two formulations are equivalent in the non-relativistic context: no relationalist will assent to either sort of Leibniz-Clarke counterfactual, and any substantivalist who endorses one sort should endorse the other. Notice that in this domain, neither sort of counterfactual has any purchase in the vacuum case: there are no material bodies to shift, so that neither sort of Leibniz-Clarke counterfactual is of much interest.

The situation changes when we shift our attention to the case of vacuum GR. The first sort of Leibniz-Clarke counterfactual remains uninteresting. But now, because of the dynamical nature of the spacetime of GR, an option opens up which was unattractive in the non-relativistic domain: we can imagine distinct possible worlds which share a geometry and a set of spacetime points, and which differ only in regard to which points have which geometrical properties.[84] Thus, given a three dimensional Riemannian geometry (Σ,g) which has a single distinguished point with maximum scalar curvature, some three dimensional substantivalists about GR will hold that it makes sense to imagine many possible worlds in which a set S of spacetime points forms a Cauchy surface whose geometry is accurately

[84] In the original presentation of the hole argument, Earman and Norton 1987, it is taken for granted that this possibility *must* be taken seriously in the non-relativistic domain. But we can always refuse to do so, by considering the geometrical properties of the spacetime of non-relativistic physics to be part of their essence. For this sort of approach, see Earman 1989, §9.6; Maudlin 1990; or Stachel 1993.

modeled by (Σ,g). There is the world in which it is *this* point has maximum curvature, and the world in which *that* point has maximum curvature; these worlds are distinct because the points of spacetime have different geometrical properties in the two worlds, although the worlds have the same geometry. A corresponding difference will be recognized by some four dimensional substantivalists about GR. We will call substantivalists who endorse the (second variety of) Leibniz-Clarke counterfactuals about GR *non-Lockean substantivalists*, to distinguish them from those who follow Locke in considering themselves to be substantivalists despite their rejection of all Leibniz-Clarke counterfactuals.

In addition to this bit of terminology, we will also want to make use of the distinction between observables and beables (see §6.5). A beable is a function on phase space which is unable to distinguish between points which represent the same physical situation; the set of beables, taken as whole, is able to distinguish between two points in phase space iff they represent distinct physical states. Similarly, an observable is a function on phase space which is unable to distinguish between points which represent situations which are observationally indistinguishable; the set of observables, taken as a whole, is able to distinguish between points iff they represent points which are observationally distinguishable. Thus, beables correspond to properties which individuate things, while observables correspond to properties which allow us to empirically distinguish between things. x and y are observationally indistinguishable iff they share all of their observable properties; they are identical iff they share all of their beable properties.

Recall that we call an interpretation of a given formulation of a gauge theory *gauge invariant* if the each gauge orbit represents a single physically possible situation. This is equivalent to saying that the beables are gauge invariant functions. This, in turn, is equivalent to saying that the beables commute with the constraints. Similarly, an interpretation is *predictable* iff the observables commute with all of the constraints. On the other hand, we call an interpretation of a given gauge system *literal* if each point in phase space represents a distinct physically possible situation—i.e., if the beables do not commute with any of the constraints. Such an interpretation is, of course, indeterministic. Note, however, that indeterministic interpretations need not be unpredictable—the observables can commute with the constraints even when the beables do not.

I can now state the conclusions which I will argue for in the next two sections. I begin in the next section by considering the possibility of giving a substantivalist interpretation of GR. Here there are four options: (i) three dimensional non-Lockean substantivalism; (ii) four dimensional non-Lockean substantivalism; (iii) three dimensional Lockean substantivalism; and (iv) four dimensional Lockean substantivalism. Both (i) and (ii) are literal interpretations of the full phase space formulation of GR; so the beables of the theory do not commute with either the vector or the scalar constraints under these interpretations. (iii) is a literal interpretation of the formulation of GR which is based on superspace, so that under this interpretation the beables commute with the vector but not with the scalar constraint. (iv) is a gauge invariant interpretation, according to which the beables of the theory commute with all of the constraints.

In §5, I turn my attention towards relationalist interpretations of GR. I argue that no relationalist interpretation can have beables which fail to commute with the vector constraint. Thus there cannot be a literal relationalist interpretation of the full phase space formulation of GR. I will argue, however, that it *is* possible to give a literal three dimensional relationalist interpretation of the partially reduced formulation of GR which is based on superspace. That is: there is a three dimensional relationalist interpretation in which the beables commute with the vector constraint, but not with the scalar constraint. I go on to argue that every four dimensional relationalist interpretation is gauge invariant—the beables of such an interpretation commute with all of the constraints.

I also argue that we can supplement any of these ontological interpretations, substantivalist or relationalist, with an account of measurement under which the theory is predictable—even if the theory is indeterministic because the beables fail to commute with some of the constraints. These conclusions are summarized in Figure 1.

Interpretative Option	Beables Commute with Vector Constraint?	Beables Commute with Scalar Constraint?	Observables Commute with Vector Constraint?	Observables Commute with Scalar Constraint?
non-Lockean 3D Substant.	No.	No.	Can, but need not.	Can, but need not.
non-Lockean 4D Substant.	No.	No.	Can, but need not.	Can, but need not.
Lockean 3D Substant.	Yes.	No.	Yes.	Can, but need not.
Lockean 4D Substant.	Yes.	Yes.	Yes.	Yes.
3D Relationalism	Yes.	No.	Yes.	Can, but need not.
4D Relationalism	Yes.	Yes.	Yes.	Yes.

Figure 1: Interpretative Options for GR

4. In this section I discuss substantivalist interpretations of GR, both Lockean and non-Lockean. For this purposes, it is helpful to begin by fixing a model (M,) of GR, and a dynamical trajectory $\tau\in\mathbf{R}x(\tau)=(p(\tau),q(\tau))$ in the full phase space, N, of the theory. We can use $x(\tau)$ to construct a model $\mathcal{M}=(\Sigma\times\mathbf{R},g)$ by letting g be the Lorentzian metric on $\Sigma\times\mathbf{R}$ which induces $q(\tau)$ and $p(\tau)$ as the first and second fundamental forms on the Cauchy surfaces $\Sigma\times\{\tau\}$. We assume that $x(\tau)$ represents (M,) in the sense that there is a diffeomorphism Φ: $M\rightarrow\Sigma\times\mathbf{R}$ such that $\Phi^*=g$.

We now fix two other dynamical trajectories, $\tau x'(\tau)$ and $\tau x''(\tau)$. We stipulate that for $\tau\leq 0$, $x(\tau)=x'(\tau)=x''(\tau)$, while for $\tau>0$, we have that $x'(\tau)$ differs from $x(\tau)$ by a gauge transformation generated by the vector constraint, and that $x''(\tau)$ differs from $x(\tau)$ by a gauge transformation generated by the scalar constraint. We go on to construct $\mathcal{M}'=(\Sigma\times\mathbf{R},g')$ and $\mathcal{M}''=(\Sigma\times\mathbf{R},g'')$ in the obvious way. We can, of course, find diffeomorphisms Φ',Φ'': $M \rightarrow \Sigma\times\mathbf{R}$ such that $g'=\Phi'^*$ and $g''=\Phi''^*$—so all three dynamical trajectories represent (M,). We can also find diffeomorphisms $d',d'':\Sigma\times\mathbf{R}\rightarrow\Sigma\times\mathbf{R}$ such that $g'=d'^*g$ and $g''=d''^*g$. Notice that since x' differs from x by a gauge transformation generated by the vector constraint, we have that d' fixes each $\Sigma\times\{\tau\}$—so d' corresponds to a re-identification of points at a given instant. Equivalently, Φ and Φ' are identification maps for the same foliation and time function on M. Meanwhile, since x" differs from x by a gauge transformation generated by the scalar constraint, we have that d" does not fix the $\Sigma\times\{\tau\}$; equivalently, Φ and Φ' are identification maps which correspond to distinct foliations on M or to distinct time functions for the same foliation. Thus, while x'(1) represents the same three geometry as x(1), we in general expect x"(1) to represent a different three geometry—either because it corresponds to a Cauchy surface from a different slicing or because it corresponds to a different Cauchy surface in the same slicing.

Let's begin by considering a non-Lockean three dimensional substantivalist interpretation of GR. Here we take the ontology of the world to consist of a substantival three dimensional space, whose points maintain their identity over time. Having established a correspondence between the points of space and the points of the three manifold Σ, a point $x=(p,q)\in N$ tells us about a possible geometry of space (and its 'time derivative'). In fact, once we fix such a correspondence, each point in N will represent a distinct possible geometry of space. Under any three dimensional substantivalist interpretation, we will recognize that two points, $x,x'\in N$, which are not related by a gauge transformation generated by the vector constraint represent distinct possible geometrical states of physical space, since they correspond to two distinct metric structures of space. Non-Lockeans, however, will claim that x and x' represent distinct possible states even when they *are* related by a gauge transformation generated by the vector constraint. That is, they will see x and x' as representing the same geometric structure on space, but insist that they nonetheless correspond to distinct states because they differ as to how the geometrical properties are distributed over the points which make up space (is it *this* point or *that* one where the geometry has maximal curvature?). Thus, under our non-Lockean three dimensional substantivalist interpretation, the three dynamical trajectories will represent three distinct possible worlds—since x(1), x'(1), and x"(1) represent three distinct possible gravitational states for the points of space. So the interpretation is both literal and indeterministic—the state represented by x(0) has multiple possible futures.

Notice that the arguments of the previous paragraph show that the beables of any three dimensional substantivalist interpretation cannot commute with the scalar constraint, since if x and x' represent distinct three geometries, they must be recognized as representing distinct possible ways that space could be at a given instant. On the other hand, the beables of a three dimensional substantivalist interpretation commute with the vector constraint iff the interpretation is Lockean—since if x and x' correspond to the same three dimensional geometry, then to say that they represent distinct states is just to say that the Leibniz-Clarke counterfactuals are true. Thus a three dimensional substantivalist interpretation is a literal interpretation of the full phase space formulation of GR if it is non-Lockean, and is a literal interpretation of the partially reduced formulation based on superspace if it is Lockean.

Of course, three dimensional interpretations of GR are quite implausible, since any such interpretation is predicated on the assumption that there is a preferred slicing of a relativistic spacetime into surfaces of simultaneity. Not surprisingly, this incompatibility with the general covariance of GR means that three dimensional interpretations are seldom taken seriously. A three dimensional substantivalist interpretation is even worse, since it underwrites unwanted notions of absolute motion. Here, as in the non-relativistic domain, the notion of absolute velocity is of no scientific use. And, while there *is* a frame-independent notion of acceleration in GR, it *cannot* be identified with acceleration relative to substantival space in any non-trivial spacetime.[85] We can avoid these unpalatable

---

[85] A trajectory is non-accelerating iff it is a geodesic of the spacetime metric. Let $\{x^\mu\}$ be the co-moving coordinates of the points of space. Look at a curve tx(t). Then a particle moving along this curve is non-accelerating relative to space iff , whereas it is a geodesic iff . Thus the two notions coincide iff all of the Christoffel symbols vanish in our coordinates. But this is impossible if the curvature is non-zero.

consequence by moving to a four dimensional substantivalism. Here our possible worlds are composed of a manifold of physical spacetime points and the set of their geometric properties (the latter can be reduced to the set of pairwise geodesic distances between points). No preferred notion of simultaneity is introduced and, since the spacetime points enjoy only a fleeting existence, the ontology cannot be used to construct notions of absolute motion other than those which derive from the metric structure of spacetime.

Non-Lockean four dimensional substantivalist interpretations are, however, still literal and hence indeterministic. Let us fix a correspondence between the points of physical spacetime and the manifold $\Sigma\times\mathbf{R}$. For a non-Lockean, each Lorentz signature metric on $\Sigma\times\mathbf{R}$ tells us about a way that geometric properties could have been assigned to the points of physical spacetime. Each metric describes a distinct possible world. Even if h and h' are metrics related by a diffeomorphism—so that they describe the same four dimensional geometry—the non-Lockean will recognize them as corresponding to distinct possible worlds, since they will differ as to their assignment of geometric properties to individual points. In particular, unless we are Lockeans, each of the solutions $x(\tau)$, $x'(\tau)$, and $x''(\tau)$ will correspond to different histories of the system. If we focus on the hypersurface S in physical spacetime which corresponds to $\Sigma\times\{1\}$, then $x(1)$, $x'(1)$, and $x''(1)$ correspond to three distinct possible assignments of geometric properties to the physical points which make up S. This is just to say that the beables of this interpretation do not commute with either of the constraints. Thus the interpretation is literal, and so indeterministic.

Lockean four dimensional substantivalist interpretations, on the other hand, are gauge invariant, and hence deterministic. Indeed, let us once again fix a correspondence between the points of physical spacetime and the points of $\Sigma\times\mathbf{R}$. Now consider the three mathematical spacetimes $\mathcal{M}$, $\mathcal{M}'$, and $\mathcal{M}''$ which correspond to our three dynamical trajectories $x(\tau)$, $x'(\tau)$, and $x''(\tau)$. Since all three instantiate the same four dimensional Lorentzian geometry, the Lockean will deny that there is any sense in saying that they represent distinct ways of assigning spatiotemporal properties to the points of spacetime. To accept this claim would be to admit that there are distinct but qualitatively identical worlds composed of the same spacetime points, which is nonsense to the Lockean substantivalist. Thus the beables of this Lockean interpretation commute with all of the constraints—i.e., the interpretation is gauge invariant. Which is to say that all of the points in a gauge orbit of GR correspond to a single possible world. This world is determined by the following process: choose $x \in N$ and look for a metric g on $\Sigma\times\mathbf{R}$ such that x can represent the geometry of a Cauchy surface in $(\Sigma\times\mathbf{R},g)$; now find a possible world, W, composed of spacetime points with a geometry modeled by $(\Sigma\times\mathbf{R},g)$; since our interpretation is Lockean, there is no other possible world with the same spacetime points and the same geometry; on the other hand, any x' which is related to x by a gauge transformation, or any g' which is related to g by a diffeomorphism would lead us to the same W.

At this point, it may seem that four dimensional Lockean substantivalism is the only viable substantivalist option for interpreting GR, since each of the other varieties of substantivalism is indeterministic. In order to counteract this perception, it is worth emphasizing that this indeterminism need not have any empirical consequences—we can supplement any substantivalist ontology with a measurement theory which makes the theory perfectly predictable. That is: we can make the observables commute with the constraints, even when the beables do not.

This is straightforward in the case of the vector constraint. Suppose that we are committed to a substantivalist interpretation of GR according to which two points $x,x' \in N$ which are related by a gauge transformation generated by the vector constraint represent distinct physical possibilities. How could we experimentally distinguish a world in which x obtained from a world in which x' obtained? Since the points are related by the vector constraint, they describe the same geometry. So the two worlds under consideration are qualitatively identical. That is, the same properties are instantiated in both worlds—the only difference lies in which individuals play which role.

The only way in which we could decide which of these two worlds we live in, is by setting out to determine the value of the physical quantity which corresponds to the function on phase space which tells us what the scalar curvature is at $y \in \Sigma$. But we could never be in a position to be able to do this, since in order to successfully carry out such a determination we would need to know *where* to go to measure the curvature. That is, we would have to know which point of physical space(time) corresponded to y. But this is just what we hope to find out by making the measurement.

This is not, of course, to say that we cannot measure the curvature at a given spacetime point. We *can* do so, and as a by-product, we discover the curvature of the point which happens to corresponds to y. It is important here to make a distinction between determining the value of a function on phase space and measuring an observable. In conducting any measurement, we determine the value of infinitely many functions on phase space. But we determine the value of a single observable. This observable is singled out from the class of functions on phase space by fact that in conducting a measurement, we don't just do something, we do something in accord with a procedure.[86] And because we are following a procedure, there is a fact of the matter about what our measurement process would have turned up had things been different.[87] That is, our procedure picks out a unique function on phase space, where the mere result of our measurement only serves to single out the class of functions which have the given value at the given point.

Suppose that we measure the curvature at a given point, which happens to correspond to $y \in \Sigma$. In doing so, we cannot have followed the procedure "go out and measure the curvature at the point corresponding to y"—since we could never be in a position to know how to carry out this instruction—and so the observable which we measure does not correspond to the quantity "the curvature at the point corresponding to y." What we actually do when we measure things is to follow procedures like "schlep the

[86] Measurement a species of intentional action. So a measurement is always a measurement under a description. See Essay 1 of Davidson 1980.
[87] 'Observ*able*' is a modal term.

equipment up to the top of the Eiffel tower, and measure the curvature there," so that what we actually measure is the observable corresponding to "the curvature at the apex of the Eiffel tower." Since we are working with vacuum GR, our instructions for measuring observables will be of the form "go out and measure the curvature in the region of maximal curvature" or "go out and measure the curvature at the boundary the region where some particular curvature scalar is zero," etc. These instructions single out observables which are gauge invariant functions on phase space. This is because the instructions refer only to qualitative properties, so that they will not produce different measurement outcomes when carried out at two qualitatively identical possible worlds. That is, these procedures single out beables which commute with the vector constraint. We measure functions on phase space which correspond to configurations of the metric tensor, rather than value of the metric tensor at a point.

This is the point that Kuchar is making when he comments that:

> the metric is not … an observable with respect to diffeomorphisms. Two metric fields, q(x) and q'(x), that differ only by the action of Diff Σ, i.e., which lie on the same orbit of $h_a(x)$, are physically indistinguishable. This is due to the fact that we have no direct way of observing the points $x \in \Sigma$.[88]

He advocates requiring the observables of GR to commute with the vector constraint. But notice that is not because he is a relationalist—he is careful to talk of the points as existent, although indistinguishable. This sort of move is available within any substantivalist interpretation of GR: one can always find a measurement theory according to which the observables commute with the vector constraint.

The situation with respect to the scalar constraint is slightly more complicated. Three dimensional substantivalists may run into trouble, depending on the sort of account of measurement that they give. If "the present volume of space" is an observable quantity, then the interpretation will not be deterministic, since this quantity does not commute with the scalar constraint—x(1) and x"(1) correspond to distinct geometries which may well have distinct volumes.

[88] Kuchar 1993a, p. 136. I have slightly altered the original notation.

There is, however, a way to get around this difficulty. An observer inhabiting the world corresponding to $x(\tau)$ can consider the spacetime which results from stacking all of the consecutive geometries of space (i.e., the spacetime represented by $(\Sigma\times\mathbf{R},g)$, where $q(\tau)$ and $p(\tau)$ are the first and second fundamental forms of $\Sigma\times\{\tau\}$). This spacetime is, of course, a derivative structure from the point of view of three dimensional substantivalism. But it may still be useful in characterizing the observables of the interpretation. In particular, $x''(1)$ corresponds to a spacelike slice in the derived spacetime, even though it does not accurately represent the geometry of space at any particular time. We can construct our measurement theory so that an observer in the world represented by $x(\tau)$ who is moving with respect to substantival space and who is carrying out measurements of curvature, measures the curvature of some spacelike slice in this fictional spacetime which, like $x''(1)$, is not space itself. Thus the observable quantities of such an interpretation may really be “the volume of the slice with such and such geometrical properties” rather than “the volume of space.”

Following this sort of route, the proponents of a three dimensional substantivalism about GR can effectively convert to four dimensional substantivalism when it comes time to face the problem of the commutation of the observables of the theory with the scalar constraint. And here all substantivalists can pick up a theme introduced by Einstein in his 1916 paper on GR. Einstein had worried for several years about his hole argument, which focused on the same puzzle that occupies us here—namely, that the general covariance of GR suggests that the theory is indeterministic.[89] In his seminal paper on GR he presents the following argument which is supposed to finesse this concern:

> All our space-time verifications invariably amount to a determination of space-time coincidences. If, for example, events consisted merely in the motion of material points, then ultimately nothing would be observable but the meetings of two or more of these points. Moreover, the results of our measurings are nothing but verifications of such meetings of the material points of our measuring instruments with other material points, coincidences between the hand of a clock and points of the clock dial, and observed point-events happen at the same place and the same time. (Einstein 1952, p. 117).

[89] See Norton 1984 and Stachel 1989 for accounts of this period, and Norton 1993 for a survey of the subsequent tumultuous history of the concept of general covariance.

As a response to the threat of indeterminism these remarks verge on operationalism. There is considerable disagreement as to whether or not it is possible to save Einstein from this fate (compare Howard 1992 with §9.7 of Earman 1989).

It should be clear, however, that Einstein's remarks are unexceptionable as a response to the accusation of unpredictability. They amount to the same point as was made concerning the vector constraint—qualitative identity implies empirical indistinguishability. In the case of GR, some sorts of substantivalists get into trouble with determinism because they are willing to countenance the existence of worlds which are qualitatively identical, but a ng assignment of geometrical properties to physical points. But there is nothing in substantivalism which commits one to the doctrine that it is possible to empirically distinguish between qualitatively identical worlds. Indeed, so long as the measurement theory is constructed using only the resources available within GR itself, it is difficult to see how one could possibly hold such a doctrine—neither the apparatus nor the objects of measurement are specifiable apart from certain properties of the curvature tensor, properties which are themselves carried along by the action of the constraints.

Now it may seem strange to maintain that the observables of GR commute with the constraints: time passes, and we can observe this. So there must be some physical quantities which are not preserved by time evolution. A fortiori, there must be some quantities which do not commute with the constraints which generate this time evolution. Here, however, it is helpful to return to Rovelli's evolving constants of motion (see §6.10). It *is* true that we can observe the passage of time—we can see the rocket move, and we can see its mass diminish. But this doesn't necessarily mean that there is an observable of the theory which corresponds to a measurable quantity, "the mass of the rocket," which changes over time, and which therefore fails to commute with the vector constraint. On Rovelli's view, what we measure are the gauge invariant quantities "the mass of the rocket as it docks" and "the mass of the rocket as it collides with the asteroid"; there is, furthermore, a natural one parameter family of such gauge invariant observables, each of which corresponds to the mass of the rocket at a given point in its journey. Such families, *evolving constants of motion*, allow us to maintain that all observables are gauge invariant, while still being able to make some sense of the notion that temporal evolution is a psychological reality. This reality finds its reflection within the formalism of our theory in the existence of a particular fibration of the set of observables into evolving constants of motion which corresponds to the time of common sense.[90]

As noted in §6.10, there is some controversy surrounding the notion of evolving constants of motion. This does not mean, however, that there is controversy concerning the possibility of a substantivalist holding that the observables of GR commute with all the constraints. Indeed, Kuchar, who is both a non-Lockean four dimensional substantivalist and a vocal critic of Rovelli's proposal (see §15 of Kuchar 1992, and pp. 135-42 of Kuchar 1993a), is himself committed to the claim that the observables of GR commute with all of the constraints (see the exchange between Kuchar and Rovelli on pp. 138-40 of Ashtekar and Stachel 1991). There is a broad consensus that GR is a predictable theory. Thus, all hands agree that the observables of GR commute with the constraints. To the best of my knowledge, no one questions that one way to make sense of this commutation is to regard the time of common sense as corresponding to a division of the set of observables into evolving constants of motion. What Kuchar and others question is: (i) on the one hand, whether this program provides a suitable framework in which to approach quantization (investigation of this question is on going; see Cosgrove 1996, Hajicek 1996, Hajicek and Isham 1996a, and Hartle 1996); (ii) and, on the other, whether this view of the nature of time is compatible with experience (see pp. 135-39 of Kuchar 1993a and pp. 266-67 of Unruh 1991; see Belot and Earman 1996b for critical discussion). The former is largely a technical question. The latter is a conceptual one, which turns upon the question of whether the family

[90] On Rovelli's view, this partition is not determined by the internal structure of the family of gravitational observables alone. Rather, it is also a function of the state of the system and our organic configuration. It is important to note that Rovelli himself doesn't attach fundamental importance to the set of evolving constants which correspond to the time of common sense—for him all such foliations of the set of observables stand on equal footing, so long as they lead to fruitful physical descriptions.

of evolving constants of motion which correspond to the time of human experience should play a fundamental role in gravitational physics.

5. I now turn to relationalist interpretations of GR. Here it is clear that the beables must commute with the vector constraint. For suppose x and x' are related by a gauge transformation generated by the vector constraint. Then they represent the *same* three dimensional geometry. At this point, substantivalists could claim that they nonetheless represent distinct situations in which different points play different roles. But relationalists do not have the resources to ground this sort of distinction. If a given spacelike slice, S, of physical spacetime is represented by x, then it must also be represented by x' since the relationalist ontology cannot countenance two possible ways that S could be which do not involve distinct geometries.

As in the substantivalist case, we find that the beables of a three dimensional interpretation do not commute with the scalar constraint. Here the ontology is just a relationalist space which endures through time. At each instant its structure is accurately described by some Riemannian three manifold. GR is the theory of the temporal evolution of this geometry. Notice that if $x(\tau)$ corresponds to a possible temporal sequence of geometries of space, then $x''(\tau)$ corresponds to a distinct such sequence. This is because, in general, the geometry represented by $x(1)$ does not appear anywhere in the sequence represented by $x''(\tau)$, so that these two dynamical trajectories must correspond to distinct possible geometrical histories of space. That is: the beables of a three dimensional relationalist interpretation of GR do not commute with the scalar constraint. Equivalently: such an interpretation is a literal interpretation of the superspace formulation of GR.

The situation changes when we move to a four dimensional relationalist interpretation, where the ontology is a four dimensional physical geometry which is accurately modeled by a solution to the EFE. Here it makes no sense to speak of $x(\tau)$ and $x''(\tau)$ as representing distinct physical possibilities, any more than it does to speak of x and x' as doing so. In both cases we have two representations of the same geometrical structure, which differ only as to which points play which roles; but a relationalist cannot contemplate the existence of two distinct situations which differ in a way corresponding to the difference between the representations. So four dimensional relationalism, like four dimensional Lockean substantivalism, is a gauge invariant interpretation in which each gauge orbit of GR corresponds to a single possible world.

From the above it follows that the observables of three dimensional relationalism must commute with the vector constraint, while the observables of four dimensional relationalism must commute with both the vector and the scalar constraints. But, as in the substantivalist case, three dimensional relationalists can help themselves to the measurement theory of their four dimensional counterparts. So GR, which is necessarily predictable under a four dimensional relationalist interpretation, can also be rendered predictable under the three dimensional doctrine.

6. I have used the question of the commutation of the beables with the constraints as a coarse-grained method for sorting interpretations of GR. This revealed six interpretative possibilities. This is several more than are usually considered—and I am sure that a more fine grained system of classification would reveal several varieties within each of my species. For my purposes, however, this classification in terms of commutation properties provides the right grain size.

The discussion up to this point is liable to leave one feeling that, ceteris paribus, either Lockean four dimensional substantivalism or four dimensional relationalism is the ideal interpretation of GR. After all, why accept an indeterministic interpretation when there is no need? I will argue in the next Chapter that all things are not equal. In particular, I will argue that one's approach to quantum gravity is partially determined by one's attitudes about the observables of the quantum theory, and specifically, by one's stance on the question as to whether or not these quantum observables are should commute with the quantum versions of the constraints. This, in turn, is influenced by one's beliefs about whether or not the beables of the classical theory commute with the constraints. Thus, on the one hand, a non-Lockean substantivalist will have different views about quantum gravity than a relationalist, so that interpretative stances concerning the classical theory will affect one's beliefs about the quantum theory. On the other hand, difficulty in making progress with the quantum theory can lead us to re-think our interpretative beliefs about the classical theory. We will see in the next Chapter that this sort of effect can even lead us to look seriously at three dimensional interpretations of GR.

Thus, I take the attitude that each of the interpretations outlined above is viable. My goal in this dissertation is to argue that there is an interesting interplay between the interpretative issues of the classical and quantum theories, rather than to argue that one or the other interpretation of the classical theory is to be preferred.

That being said, it is nonetheless tempting to make some comparisons, based solely upon considerations internal to GR itself.

As I have stressed several times, there is nothing within classical GR to justify the heterodox step of breaking the theory's general covariance by adopting a three dimensional interpretation. Minkowski's prediction that space and time would lose their independent existence has been born out by the subsequent history of physics. Thus, upto quantum issues, four dimensional interpretations are preferable to three dimensional interpretations. *Within* the category of three dimensional interpretations, we will presumably prefer three dimensional relationalism to Lockean three dimensional substantivalism—the beables of the two interpretations have identical commutation properties, but only the latter is committed to a hierarchy of spurious concepts of absolute motion. From the perspective of the classical theory, non-Lockean three dimensional substantivalism is likely to appear even more unfortunate, since its beables do not commute with any of the constraints.

Comparisons within the class of four dimensional interpretations lead us into the domain of the hole argument literature. Here the dialectic has been as follows. In their original hole argument paper, Earman and Norton 1987 establish a connection between substantivalism about GR and indeterminism, by arguing, in effect, that four dimensional non-Lockean substantivalism is indeterministic. Their argument—which is a re-tooled version of Einstein's original argument—is very similar in spirit to, and was the inspiration for, the discussion of §4 above.

Earman and Norton's paper spawned a multitude of responses, most of which are attempts to show that substantivalism can be rescued from the clutches of indeterminism. These responses fall into two broad categories. The first group was discussed in §6.2. It is comprised of discussions of the hole argument which rely on a formal approach to determinism (see Healey 1995, Leeds 1995, Mundy 1992, Rynasiewicz 1992, Rynasiewicz 1994, Rynasiewicz 1996b, and Wilson 1993). I believe that the discussion of Chapter 6 shows that these responses are untenable. The fact is that every gauge theory admits interpretations under which it is indeterministic. Since GR is a gauge theory, the program which these authors attempt to carry out—proving that GR is deterministic, independently questions of interpretation—is patently impossible.

The second group of replies recognizes the connection between determinism and ontology (see Bartels 1994, Brighouse 1994, Brighouse 1995, Butterfield 1987, Butterfield 1989b, Butterfield 1989a, Hoefer 1996, Maidens 1993, Maudlin 1989, Maudlin 1990, Stachel 1989, and Stachel 1993). Here it is allowed that Earman and Norton have succeeded in drawing a connection between indeterminism and a certain brand of substantivalism. Furthermore, it is taken for granted that this indeterminism redounds to the discredit of this variety of substantivalism, even should the latter be supplemented with a predictable measurement theory.[91] But, it is argued, the sort of non-Lockean substantivalism which Earman and Norton consider is only one of many possible substantivalisms—and it is the

[91] "Determinism may fail, but if it fails it should fail for a reason of physics, not because of a commitment to substantival properties which can be eradicated without affecting the empirical content of the theory," Earman and Norton 1987, p. 524. This thesis strikes many as questionable. See, for instance, critical discussion of Hoefer and Cartwright 1993 and Hoefer 1996.

most naive one at that. The heart of this second sort of defense against the hole argument is an elaboration of a Lockean form of four dimensional substantivalism which is gauge invariant, and hence deterministic. Thus, substantivalism is supposed to be rescued.

At this point, a great deal of effort is expended in formulating and defending Lockean substantivalisms. Here there are two approaches. The first involves an attempt to ground Lockeanism in a general account of modality, while treating spacetime points as run of the mill existents (Bartels, Brighouse, Butterfield, and Maudlin). The second involves simply stipulating that spacetime points are queer existents whose modal semantics is surprising in a number of ways, one of which is that the Leibniz-Clarke counterfactuals come out false (Hoefer, Maidens, and Stachel). And now the battle is joined—see Earman 1989, Norton 1989, Belot 1995b, and Norton 1989, Belot 1995afor criticisms of some of the above approaches.

Thus, the debate over the hole argument plays itself out as a debate over the viability of Lockean substantivalism. It is taken for granted that if Lockean substantivalism is preferable to non-Lockean substantivalism—the only question is whether the former exists.

My own view is that the priorities of the debate are inverted. I do not think that it is difficult to advance a viable Lockean substantivalism. I do not, indeed, believe that this has yet been done within the modal approach to Lockeanism.[92] But, as discussed in §3.7, I believe that the approach of Maidens is perfectly satisfactory—one may view the hole argument as teaching us that spacetime points have to be handled with semantic care. This involves *stipulating* that the Leibniz-Clarke counterfactuals are false, rather than deriving this result from some more general framework. As such, it is more like theft than honest toil. For this reason, it is likely to seem repugnant to those who are not already deeply committed to substantivalism.

I do not, then, question the existence of Lockean substantivalism. But I do question its attractiveness.

[92] I relegate the justification of this claim to the appendix, since it would take us too far afield at this point.

In moving to a Lockean position, substantivalists give up a lot. In shifting from three to four dimensions, substantivalists have already moved away from the intuition that space is substance-like. Non-Lockean substantivalists can maintain some contact with their roots by affirming the Leibniz-Clarke counterfactuals. But Lockean substantivalists are committed to a position which is not only far-removed from the original motivation, but which is also correspondingly more difficult to understand. What exactly does it mean to say that spacetime points exist, if they do not endure over time or support counterfactuals? Here Lockean substantivalists are likely to entrench themselves, and to insist that they know what existence is, and they know that spacetime points have it—and if their opponents are unlike them in either respect, then so much the worse for them. This is fair enough. But they must realize that this sort of dogmatic approach is unlikely to win many new converts.

In addition to this loss in intelligibility, there is a loss in distinctiveness: Lockean substantivalism is very like relationalism. Indeed, the middle two rows of Figure 1 are identical to the last two rows. It is sometimes held not only that four dimensional relationalism and four dimensional substantivalism about GR are so similar to one another that the substantival-relational debate is on the verge of collapse, but also that relationalism is clearly the loser in this rapprochement. The claim here is that relationalists have ceded far more of their original territory (see §§6-7 of Maudlin 1993). It should already be clear, on strictly intratheoretic grounds, that I think that this view is off the mark. It is by no means clear that it is relationalists who have given up more in adapting their view to GR—especially not if it is Lockean substantivalists that they are being contrasted with.

When it comes to intertheoretic considerations, I think that Lockean substantivalism is the clear loser.[93] In the next Chapter, I will argue that the interpretative interaction between classical and quantum gravity breathes new life into the substantival-relational debate—so long as it is non-Lockean substantivalism which is at stake. My argument in favor of this claim depends crucially on the difference between the beables of the respective

[93] Although I focus on the interplay between classical and quantum gravity, I do not mean to rule out the existence of other intertheoretic—or even nontheoretic—influences.

interpretations. But when it comes to the opposition between relationalism and Lockean substantivalism, I can think of no relevant difference between the two doctrines which would lead to any interesting interplay between serious physics and the (increasingly metaphysical) issue between relationalism and Lockean substantivalism.

Thus I do not consider the adoption of Lockean substantivalism to be an attractive response to the hole argument. It is true, of course, that Lockean four dimensional substantivalism is a deterministic interpretation of GR. But this will seem to be an advantage only so long as the hole argument is seen as presenting a threat to substantivalism. I would claim, however, that the hole argument presents substantivalists with an opportunity rather than a threat. The kernel of the hole argument is the observation that non-Lockean substantivalism is a literal interpretation of GR, formulated as a gauge theory. Cast in one way, this observation is threatening—“under a non-Lockean substantivalist interpretation, GR is an indeterministic theory.” Cast in another, it becomes an opportunity—“the beables of non-Lockean substantivalism are unlike those of relationalism in failing to commute with the constraints of GR.” Here we have a substantive difference between substantivalism and relationalism that has the prospect of being relevant to the search for a quantum theory of gravity. I propose that, rather than taking the conservative course of advocating a safe brand of substantivalism which is little more than a variation on relationalism, substantivalists should be eager to take the bold course and to attempt to exploit the type of beables which are only available within their account. The hope here is that this will give them a decisive advantage over relationalism when the dust surrounding quantum gravity finally settles. That is, they should seek to entitle themselves to the advantage that they have always claimed over their relationalist rivals—vindication by physics rather than metaphysics.

Lockeanism is not the savior of substantivalism: rather, it is a pallid imitation of relationalism which should be interest only to those substantivalists who are too cowardly to wager that quantum gravity will carry full-blown non-Lockean substantivalism to a decisive victory over relationalism.

CHAPTER 8
Quantization and Interpretation

1. Introduction

Quantization is a technique for producing quantum versions of classical theories. Here we focus on canonical quantization, where the initial input is a classical theory which may either be a true Hamiltonian theory or a gauge theory, and the output is a quantum theory which—one hopes—has the original classical theory as its classical limit. It is important to emphasize that neither quantization, nor the taking of a classical limit is a straightforward process. There is no universally applicable algorithm for either sort of process. The details greatly from case to case.

Indeed, most classical systems have many possible quantizations. This is, notoriously, true for of all field theories, and is intimately related to puzzling features of quantum field theory such as the Unruh effect (see example 3 below). But this problem is endemic even among classical systems with finitely many degrees of freedoms (see example 2 below)—although this fact is usually glossed over in even the best discussions of the philosophical foundations of quantum mechanics.

Similarly, there is no universally applicable technique for taking a classical limit. In some cases, one will be satisfied with a WKB approximation, in others a result along the lines of Ehrenfest's theorem will be required, in still other cases one will want to see certain sorts of qualitative behavior emerge as certain parameters are varied (typically, as $h \to 0$). Even when a satisfactory classical limit exists, it is by no means a foregone conclusion that this limit will be the classical system upon which the quantum system was based. In recent years there has been a great deal of discussion of this possibility in the particular case where one is dealing with a quantization of a chaotic classical system with finitely many degrees of freedom. The question of the correct notion of classical limit for such systems has been at the center of this debate (see Belot and Earman 1996a for a critical discussion of this literature).

Thus there is considerable play at either end whenever we attempt to set up a correspondence between classical and quantum systems. This freedom in bridging the classical-quantum divide is closely related to an interplay which exists between interpretative issues at the classical and quantum levels.

On the one hand, interpretative issues can guide one in choosing a quantization or a classical limit. This is particularly clear in the case of quantization. Ultimately, one hopes that empirical considerations will resolve the ambiguities involved in quantization. That is, one hopes that experiment will be able to determine which of the available quantizations of a given classical system is an adequate representation of actual physical systems (this happens with the Aharonov-Bohm effect; see §6.8 and example 2 below). But we sometimes find ourselves in a situation where we must choose a quantization, or an approach to quantization, without guidance from the tribunal of experience. This happens, for instance, in quantum field theory on a curved background, where there are no available empirical data. Quantum gravity provides an even more extreme case. Many physicists expect that the theory will produce novel predictions only for situations involving magnitudes on the Planck scale.[94] Since the Planck energy, at 10^{18} GeV, is *many* orders of magnitude above what can be expected to be empirically accessible in the foreseeable future, there is not even the prospect of such data becoming available. Here empirical data seriously underdetermines the choice of a quantum theory or an approach to quantization. And yet, as in Pascal's wager, one feels that one choose. Here there are probably a number of different epistemic elements which can be deployed to fill the logical gap between the classical and the quantum. One such element is the set of one's interpretative beliefs about the content of the classical theory. This can shape one's preferences when confronted with competing quantizations, or influence one's choice of quantization scheme.

---

[94] This is the scale upon which gravitational interactions between elementary particles should be discernible. It remains an open question whether quantum gravity will make predictions which can be tested at the cosmological level.

Conversely, the choice of a quantization or classical limit can have interpretative consequences at the classical limit. It has, for instance been argued that the anomalous nature of the classical limit in the presence of chaos provides support for the interpretative claim that classical mechanics should be understood as a theory of ensembles rather than individual systems (see Ballentine 1996). Furthermore, if the tribunal of experience rules in favor of Quantization A over Quantization B, we will tend to look askance at interpretative approaches to the classical theory which tell us that B is preferable to A (see examples 2, 4, and 5 below). Finally, interpretative moves at the quantum level can make us rethink our interpretations of the classical theory (see example 5 below).

This Chapter is dedicated to exploring these sorts of interpretative interrelations between the classical and the quantum. I have two goals in mind here. The first is by now familiar: I want to demonstrate that the substantival-relational debate is a live and interesting topic by showing that there is a relation of reciprocal dependence between ontological questions concerning the spacetime of GR and proposed solutions to the conceptual problems of quantum gravity. My second reason for wanting to examine the interpretative interrelation between the classical and the quantum was briefly discussed in §1.4 and will be revisited in the Afterword: I believe that the examination of intertheoretic relations is a crucial component of the project of understanding what our theories tell us about the world.

The Chapter is structured as follows. In §2 I present a quantization schema—this is a general program whose details can be filled-in in many ways. It is by no means the most general schema, but it is general enough to cover the cases that I will be interested in in this C quantization involving finite dimensional classical systems and linear fields. These illustrate the quantization schema sketched in §2, and provide paradigm examples of the interpretative interrelation between the classical and the quantum

I then turn to my central example: the problem of quantizing gravity. The search for a quantum theory of gravity is an on-going research program. In this Chapter I will focus on one branch of this program—canonical quantum gravity. Here the goal is to construct a quantization based upon the sort of Hamiltonian formulation of GR which was considered in the previous Chapter. This is a formidable task for a number of reasons. First of all, GR is a nonlinear field theory—and the quantization of such theories requires some quite sophisticated mathematics. Second, the gravitational field is unlike other physical fields in that it cannot be modeled as propagating against a fixed spatiotemporal background. One therefore expects that the search for a quantum theory of gravity will require the development of new techniques, over and above those required for other quantum field theories. Finally, one finds that even if one imagines away all of these technical problems, there remain some outstanding conceptual problems. In particular, the role of time in the quantum theory is extremely difficult to understand—and this difficulty appears to be a direct result out of the general covariance of GR.

All of this means that gravity provides a fine example of a situation where there is considerable uncertainty about how to bridge the gulf between the classical and the quantum. And, as mentioned above, one does not expect any immediate help from empirical considerations in reducing this freedom. Thus it is not surprising that we find that there is a very rich interplay between interpretative issues of classical and quantum theories of gravity.

After sketching, in §4 the most straightforward route to canonical quantum gravity, and the severe conceptual problems that it throws up, I turn in §5 to two proposals for resolving these problems, and examine the ways in which these proposals are driven by, and drive, interpretative agendas at the classical level.

2. Quantization.

In this section I present a quantization schema. It is a set of rules which leads one from a classical system—which may be either a Hamiltonian system or a gauge theory—to a quantum system. We assume that the classical system is based upon a symplectic geometry, (M,ω). If the classical system is a Hamiltonian system, then it is of the form

(M,ω,H); if it is a gauge system, if is of the form (N,σ,H), where (N,σ) is a first class constraint manifold in (M,ω). The quantum system consists of a Hilbert space together with a set of self-adjoint operators (the observables), and a Hamiltonian operator which governs the dynamics via the Schrödinger equation. The quantum system is the quantization of the classical system in the minimal sense that there is a bijective correspondence between some subset of the set of observables of the classical system and the set of observables of the quantum system such that (where [.,.] is the commutator, and {.,.} is the Poisson bracket).[95] Two provisos are in order. First: requiring the algebra of quantum operators to have the same structure as a subalgebra of the classical Poisson is a quite weak condition, and by no means guarantees that the original classical system is the classical limit of its quantization.[96] Second: we are concerned with formalism rather than interpretation at this stage, so that nothing should be read into the use of the term 'observable'—in example 5 below, we will return to this issue, and ask whether the quantum observables should correspond to classical observables or to classical beables.

The quantization schema should not be thought of as being an *algorithm* for quantizing classical theories. That is, the schema is not a black box which takes classical theories as input and spits out quantizations. Rather it is a set of instructions, each of which is vague and open ended. As such, there is typically room for, and sometimes the necessity of, further input at each step in the schema. Thus there is considerable opportunity for conceptual, technical, and empirical considerations to influence the output of the process.

The schema which I present is only one of many possible approaches to quantization. It should probably be thought of as being of middling abstraction. On the one hand, the approach is far more abstract that the 'put hats on the p's and q's' approach of naive treatments of quantum mechanics. This is a good thing: it is only by adopting a certain

[95] This requirement insures that the appropriate uncertainty relations will hold in the quantum theory.
[96] It is possible, for instance, to find chaotic classical theories hose quantizations are periodic. There is a sense in which such a classical system is not the classical limit of its quantization. See Belot and Earman 1996a.

level of abstraction that the freedom available in quantization is made explicit. On the other hand, it is not nearly so abstract as algebraic approach, since it involves a choice of a concrete representation of a set of an algebra of observables as a set of operators on a Hilbert space.[97] Again, this is appropriate for our purposes, since there are difficulties in a a spatiotemporal background geometry.[98]

The schema outlined below is more or less the standard schema for those working on canonical quantum gravity. It is a direct descendent of the program of Dirac 1964, and is only a slight variation on those of Ashtekar 1995 and Kuchar 1993a. There is fairly wide agreement on the shape of the schema, and on the input required at each stage. There is, however, considerable disagreement concerning the degree to which the stages can be treated independently of one another. Ashtekar regards the steps as being more or less independent, so that the program can be carried to completion by facing each of the steps in sequence. Kuchar regards the steps as being interconnected in a Kafkaesque tangle, so that one can never be sure of having passed through any of the stages until they have all been completed (pp. 147-48 of Kuchar 1993a). This issue will only be resolved if and when there is a canonical theory of quantum gravity. In any case, my quantization schema is a seven step program, and, whatever the logical relation between the steps, it is at least possible to state them as a sequence.

---

[97] Under the algebraic approach, the basic object of investigation is the algebra of quantum observables. One considers this as an abstract algebra, rather than as concrete algebra of operators on a Hilbert space. One can talk about algebraic states—which are just linear functional on the algebra. These should be thought of as returning the expectation value of the observable when the system is in the given state.

One can go on to consider concrete representations of the algebra, and to identify a subset of the set of states with vectors (or density operators) of the associated Hilbert space. It is not, however, possible to view all of the algebraic states as states associated with any given Hilbert space representation of the algebra. Nor, in the field-theoretic case, are all concrete representations unitarily isomorphic, even if we require them to be irreducible. These facts will play an important role in example 3 below.

See Haag 1992 for the details of the algebraic approach, or Saunders 1988 for an introduction aimed at philosophers.

[98] This difficulty stems from the fact that the standard formulations of the algebraic approach presuppose the existence of a fixed causal structure of spacetime. See §VII.3.4 of Haag 1992 and §1.2.2 of Ashtekar 1995.

In presenting these seven steps, I will attempt to make clear what sort of input is required at each stage. Occasionally, this will require calling attention to some of the purely mathematical difficulties associated with the construction of a quantum field theory. In general, however, I will emphasizes the big picture by glossing over the technical difficulties. It is safe to assume that I am sweeping nontrivial mathematical problems under the rug at every stage of this quantization program, as well as in the discussion of the following sections.

For the reader approaching these issues for the first time, it may not be clear why it is helpful to analyze quantization in this way. Such readers are requested to be patient—it will become apparent in subsequent sections that this approach provides a helpful framework in which to formulate and discuss the difficulties of quantum gravity.

Step (1) Choose a classical formulation. Our schema requires a gauge theory or a Hamiltonian system as its initial input. There may be considerable freedom in choosing this input—since there will be several such formulations of any given theory. In the case of GR, for instance, there are many possibilities. One could select any of the formulations considered in Chapter 5: the gauge theory based on the full phase space, the partially reduced gauge theory based on superspace, or the Hamiltonian system based on the reduced phase space of equivalence classes of diffeomorphically related models of GR. One could also opt for a Hamiltonian (or gauge theoretic) formulation of the theory which arises from one of the above via (partial) gauge fixation. Finally, it is important to note that there exist formulations of GR as a gauge theory which are completely different from those considered in Chapter 5. The most important of these is the one due to Ashtekar, where the configuration space is the space of connections rather than the space of metrics. This leads to a gauge theory with the same phase space as a Yang-Mills theory, but with additional constraints (see Ashtekar 1995). One can again construct several more formulations of GR by performing various sorts of reduction or gauge fixation on the Ashtekar formulation.

Step (2) Choose fundamental variables. It is well known that it is impossible to find a consistent quantum theory in which there is a quantum observable corresponding to every classical observable (van Hove's theorem; see, e.g., §5.4 of Abraham and Marsden 1978). At this step we choose the set of *fundamental variables*, a subset of functions on the phase space (M,ω) which is: (i) large enough to provide a set of coordinates; (ii) small enough that there will be a quantum theory whose observables correspond to the fundamental variables. Note that (i) allows us to approximate arbitrary functions on phase space by polynomials in the fundamental variables. If we are working with a gauge system, we further suppose that the constraints can be expressed nicely in terms of the fundamental variables.

Here there are typically a number of choices: in ordinary quantum mechanics, for instance, one can work with either position and momentum, q and p, or with the Bargmann variables, and .

The set of fundamental variables has an algebraic structure which is determined by the structure of the classical phase space. Part of this structure is determined by the (pre)symplectic structure of the phase space, since this determines the Poisson bracket relations between the fundamental variables. There may, however, be further structure which arises out of the global topological structure of the phase space (see §2.4 of Ashtekar 1995). We will denote the algebra of the fundamental variables by $\mathcal{A}$.

Step (3) represent the algebra $\mathcal{A}$. At this stage we look for a vector space, V, together with an irreducible representation $\rho: \mathcal{A} \to L(V,V)$ of the algebra of fundamental variables as a set of linear operators on V.[99] ρ provides the correspondence between classical and quantum observables which is the core of quantization. Notice that at this stage, V is just a vector space—it does not necessarily have an inner product, and thus may not be a Hilbert space.

[99] ρ is a homomorphism; i.e., it preserves the algebraic structure of $\mathcal{A}$. ρ is irreducible if there is no subspace of V which is invariant under all of the operators in its range.

Here there will be many possibilities. In ordinary quantum mechanics, the choice of V corresponds to the choice of 'representation,' in the sense of, for instance, the position representation as opposed to the momentum representation. The vector space associated with the position representation is $L^2(Q,d\mu)$; here one follows Schrödinger in writing and . This is quite distinct from the Fock space which one associates with the Bargmann variables.

In the quantum mechanical case—where the system being quantized has finitely many degrees of freedom—the Stone-von Neumann theorem tells us that any two irreducible representations of an algebra as an algebra of self adjoint operators on a Hilbert space are unitarily equivalent. This means that any two such representations will lead to identical predictions—so that in the finite dimensional case the freedom available in this step is of no physical significance. The situation changes radically in the infinite dimensional case, where there are always infinitely many unitarily inequivalent irreducible representations of a given algebra (see Wald 1994). Thus the choice of representation of the algebra of fundamental variables has physical content whenever we are dealing with a field theory.

Step (4) Impose constraints on set of states. Up until now, our approach has been neutral between the gauge theoretic case, and the strictly Hamiltonian case. At this point, we have to implement quantum versions of the constraints. This step is, of course, omitted if one is quantizing a Hamiltonian system.

We have assumed that the classical constraints can be expressed in closed form in terms of the fundamental variables. We now fix one such expression, $C_a(x_i)$, for each classical constraint. At this point, we write down operators which formally correspond to the classical constraints. We then restrict our attention to the subspace of V which is annihilated by the constraints. This is called the *space of physical states*:

.

The motivation for this condition is that we want the state vectors to be gauge invariant (see §13.3.5 of Henneaux and Teitelboim 1992 for another justification). To see how this works, let us return to the system of example 5.4 (recall that this example is typical of the local geometry of finite dimensional gauge systems). Here we have a symplectic manifold (T^*Q,ω) with canonical coordinates $(p_i;q^i)$, and a gauge theory (N,σ,H), where the constraint submanifold $(N,\sigma)\subset(M,\omega)$ is determined by the constraint $p^1=0$. We take the canonical coordinates as our fundamental variables, and employ the ordinary Schrödinger representation of quantum mechanics so that V is the space of square integrable functions on the configuration space Q, we have and . The quantum constraint is just . This means that V_{Phys} is subspace of V consisting of ψ for which

.

That is, V_{Phys} is the space of wave functions on Q which are independent of the gauge degree of freedom, q_1. Below we will see that the imposition of the quantum constraint for a parameterized free Newtonian particle leads directly to the ordinary Schrödinger equation, while the imposition of the quantum constraint for a free relativistic particle leads to the Klein-Gordon equation (see example 1).

There is considerable freedom available at this stage of the construction. There is, first of all, freedom in the choice of expressions for the classical constraints in terms of the fundamental variables. Then there is freedom in translating these classical expressions into expressions in terms of the operators corresponding to the fundamental variables. In particular, because the classical quantities commute with one another under multiplication whereas the operators on V in general do not, one has to make factor ordering choices at the quantum level.[100] This gives one considerable freedom. Some of this freedom is removed by the desire to avoid introducing anomalies: since for first class constraints we have $\{C_i,C_j\}=0$ at the classical level, we would also like to have =0 at the quantum level; but this

[100] We do not expect our quantum operators to commute with one another, so and must be viewed as being distinct ways to put hats on the classical expression 'ab.'

will fail for some choices of factor ordering (see §13.3.2 of Henneaux and Teitelboim 1992).

Step (5) Inner product. One now imposes an inner product μ on V_{Phys}, such that: (i) $\mathcal{H}=(V_{Phys}, \mu)$ is a Hilbert space; (ii) the operators on $\mathcal{H}$ representing the algebra of classical fundamental variables are self-adjoint. This does not ordinarily pose much of a problem in ordinary quantum mechanics (see §5.4 of Abraham and Marsden 1978). It is, however, often quite difficult to find a physically appropriate inner product for a quantum field theory. Indeed, selection of a physical inner product has turned out to be one of the most difficult tasks facing the program of canonical quantum gravity.

Step (6) Observables. So far we have a Hilbert space and a set of operators. The latter, however, may not be the observables of our quantum theory. They correspond to the classical fundamental variables. But, the fundamental variables of a gauge theory will not in general be gauge invariant (think of the metric and extrinsic curvature in GR). Thus the operators which correspond to the fundamental variables of a gauge theory will not in general represent measurable physical quantities. In this case, it is necessary to select a further subset of classical observables which we then elevate to quantum observables by finding self adjoint operators on $\mathcal{H}$ which satisfy the appropriate algebraic relations with our other observables. Note that there is, as in step 4, considerable freedom in passing from an expression for a classical observable to an expression for a quantum observable.

Here there is also a major decision to be made concerning the relation between the observables and the constraints. The standard approach, due to Dirac 1964, requires that the quantum observables commute with the quantum constraints. This is just the quantum analog of the classical requirement that the observables commute with the first class constraints, which is necessary in order for the theory to be predictable. It is generally accepted that the Dirac approach is the correct one—especially when the theory has a non-zero Hamiltonian and is not time reparameterization invariant. There is, however, some dissent in the case where the Hamiltonian is zero: in this case, some of the classical constraints generate time evolution; this means that the role of time in the quantum theory is very mysterious; it is sometimes suggested that this difficulty can be avoided if the quantum observables fail to commute with the constraints. This suggestion will be discussed in detail below in connection with the problem of time in quantum gravity (see example 5).

Step (7) Dynamics. We now have a Hilbert space and a set of observables. All that we are missing is an equation governing the dynamics of our theory. At this stage, one chooses an expression $H(x_i)$ for the classical Hamiltonian in terms of the fundamental variables, and then looks for an operator which satisfies the appropriate algebraic conditions (it must, for instance, commute with the quantum constraints, since the classical Hamiltonian commutes with the classical constraints). As usual, there is considerable choice at this stage.

We now stipulate that the dynamics of our theory is determined by Schrödinger's equation,

.

This completes the construction of our quantum theory.

Notice, however, that if we have a classical theory with a zero Hamiltonian, then we will have ψ=0 (since H is just a sum of the constraints, and the quantum constraints annihilate states). Thus Schrödinger's equation *appears* to tell us that there is no change in a quantum theory based upon such a classical theory. This is the core of the problem of time in quantum gravity.

3. In this section I present three examples of quantization. The first is meant to give a feel for how the imposition of quantum constraints proceeds, and provides a helpful background for the problem of time in quantum gravity. The second and third examples are intended to show how the freedom available in choosing a quantization of a classical system: (i) is physically relevant; and (ii) provides a fulcrum for the interaction of interpretative issues between theories.

Examples.

(1) Free particles. Consider a free Newtonian particle of mass m in a one dimensional substantival space. We can, as in example 5.2, model this system using a Hamiltonian system, (T^*Q,ω,H), where Q={x}=**R**, ω is the canonical symplectic form, and $H=p^2$ (where p is the momentum canonically conjugate to x). We can apply the usual naive technique of quantization, and end up with a Hilbert space L^2(**R**), and a Schrödinger equation: .

What happens if we parameterize before quantizing? In this case, we work with the constraint submanifold of ($T^*(Q\times\mathbf{R})$, $\omega - dt\wedge dp_t$) which is given by $H'p_t+H=0$. Since the Hamiltonian for our constrained system is just the constraint, H', our gauge theory is time reparameterization invariant (see §5.9). As discussed above in connection with step 7 of the quantization schema, this means that the imposition of the Schrödinger equation, is dynamically trivial, since our quantum states are subject to the constraint 'ψ=0. Does this mean that there is no change or time in our quantum theory? No. For notice that the constraint 'ψ=0 is equivalent to:

,

which is, of course, just the ordinary Schrödinger equation of the quantization of the unparameterized system.

Thus it is relatively straightforward to recover the dynamical picture of ordinary quantum mechanics from the prima facie timeless quantum system that results from quantizing a parameterized free particle in Newtonian spacetime. The situation is somewhat more complex when we look at the quantization of a free particle in Minkowski spacetime, (M,η_{ab}). We consider the cotangent bundle $T^*M=(p_a;x^a)$ with the canonical symplectic structure, ω. The motion of the particle is modeled by the constraint surface in (T^*M,ω) which is given by the constraint/Hamiltonian:

=0.

Again, we have a time-reparameterization invariant theory, so that the Schrödinger equation of step 7 is dynamically empty. In this case, however, we cannot expect the quantum constraint ψ=0 to be a Schrödinger equation: in the parameterized Newtonian case the constraint was linear in the momentum canonically conjugate to time, so that the quantum constraint could be solved for the time derivative of ψ; in the relativistic case, the constraint is quadratic in all of the momenta, including the momentum canonically conjugate to time, so that the quantum constraint is not a Schrödinger equation.

What, then, is the physical content of the quantum constraint for this system? Writing, as usual, , we see that ψ=0 becomes:

.

This is, of course, just the Klein-Gordon equation of example 5.3. Thus our theory of a free relativistic particle is just the theory of complex valued solutions of the Klein-Gordon equation in Minkowski spacetime.

If we now fix an inertial frame, we know from example 5.3 that we can write down a Hamiltonian system (T*Q,ω,H) which describes this theory, where Q is now the space of instantaneous configurations of the complex Klein-Gordon field. Here the symplectic structure is given by :

,

where the surface of simultaneity, Σ, and the time derivatives are induced by our choice of inertial frame. ω is independent of the choice of Σ. We now restrict our attention to the space $\mathcal{H}$⊂T*Q of initial data which give rise to solutions of the Klein-Gordon equation which oscillate with positive frequency relative to the time of an arbitrary inertial observer. Then iω(·,·) is an inner product on $\mathcal{H}$ which makes it into a Hilbert space. This inner product is independent of Σ, and is thus conserved in time. It is therefore appropriate for the calculation of probabilities in a quantum theory. Furthermore, we can use the time coordinate t of our chosen inertial frame to deparameterize the system by transforming the constraint equation =0 into a form analogous to that of the Newtonian constraint $H'p_t+H=0$. Relative to this choice of inertial frame, the quantum constraint becomes a Schrödinger equation governing the evolution of the state vector in time. This Schrödinger equation turns out to be Poincaré covariant, so that the resulting quantum theory is independent of the choice of inertial frame. Thus, the deparameterized version is equivalent to the original Klein-Gordon version.

This approach is extremely difficult to generalize to the case of a free particle in an arbitrary general relativistic spacetime (see Kuchar 1991). The construction of the conserved inner product, and the proof that the Schrödinger equation is covariant both require that the spacetime be stationary (i.e., that there be a one parameter family of timelike isometries). In general, of course, this property does not hold, and it is impossible to construct an appropriate Schrödinger equation and inner product for a free quantum particle. We will see that this fact has important interpretative consequences for quantum field theory on a fixed spacetime background and for quantum gravity (see examples 3 and §4 below).

(2) Geometric quantization and the Aharonov-Bohm effect. The ambiguity inherent in quantization is hidden in textbook presentations of quantum mechanics. There one is presented with a Hamiltonian system (T^*Q,ω,H), typically with $Q=\mathbf{R}^n$. One then selects canonical coordinates, $(p_i;q^i)$, and proceeds to quantize by using the Schrödinger representation to elevate the Poisson algebra relations between the canonical coordinates to commutation relations between the operators and on $L^2(Q,dx)$. The Stone-von Neumann theorem guarantees that any other irreducible representation of the classical algebra as an algebra of self adjoint operators on a Hilbert space will be unitarily equivalent to that of the Schrödinger representation (see, e.g., 2.2 of Wald 1994). The only ambiguity lies in selecting a complete set of quantum observables, and in factor ordering problems with the observables and the Hamiltonian. But, typically, symmetry principles will take care of these problems (see, e.g., §5.4 of Abraham and Marsden 1978).

There are a couple of obvious questions. To what extent is the above construction independent of the original choice of canonical coordinates? Is it possible to find a construction which does not rely, as the Schrödinger representation does, on the classical phase space having the structure of a cotangent bundle?[101] Both of these questions nudge one in the same direction. What one would ultimately like is a *geometric* approach to quantization, which led directly from the symplectic geometry of the classical phase space to a quantum theory without relying upon a choice of canonical coordinates. Such a theory would allow one to construct quantizations of systems, like classical spin systems, whose phase spaces do not have the structure of a cotangent bundle. It would also allow one to analyze the degree to which the geometric structure of the classical phase space singles out the structure of the quantum theory—and hence to isolate the extent to which one needs further input, such as a choice of a preferred system of canonical coordinates.

This is the motivation for the research program of geometric quantization (see §5.4 of Abraham and Marsden 1978 or §34 of Guillemin and Sternberg 1984 for an outline; see Puta 1993, Sniatycki 1980, or Woodhouse 1980 for the details). The upshot of this program is that one can proceed from a Hamiltonian system (M,ω,H) to a quantum theory in two steps.[102]

(i) Prequantization. The first step is to find a line bundle, B, over M with a connection, ∇, whose curvature is ω.[103] This bundle with connection is called a *pre-quantum bundle*. It is not always possible to construct a prequantum bundle for a given (M,ω).[104] But when it is,

---

[101] The cotangent bundle structure sorts the fundamental variables into p's and q's.

[102] Geometric quantization can be generalized to apply to gauge theories. See Ashtekar and Stillerman 1986.

[103] B is constructed by attaching a copy of **C** to each point of M. Locally, B just looks like $M\times\mathbf{C}$, in the sense that every point $x\in M$ has a neighborhood U in B of the form $V\times\mathbf{C}$, where V is a neighborhood of x in M. But globally, B may differ from the *product bundle* $M\times\mathbf{C}$. For a discussion of connections on bundles and their curvature, see, e.g., §V.B of Choquet-Bruhat, DeWitt-Morrette, and Dillard-Bleick 1982 or Chapter 7 of Nash and Sen 1983.

[104] A necessary and sufficient condition for this to be possible is that the integral of ω over any closed surface in M should be an integral multiple of 2π. This condition is intimately related to the quantization of spin and charge.

there are, in general, many possibilities: the inequivalent prequantum bundles for (M,ω) are parameterized by the cohomolgy group $H^1(M,U(1))$. It follows that there will be inequivalent prequantum bundles over (M,ω) iff M is non-simply connected.[105]

In the special case where our phase space is M=T*Q with the canonical symplectic form ω, there is an obvious choice for a prequantum bundle (B,∇): we take B to be the product bundle M×**C**, and let ∇ be the connection on B whose potential is just the canonical symplectic potential $\theta=p_a dq^a$ (θ is called the symplectic potential since ω=dθ). We will call this the *canonical prequantum bundle*.

If we now fix a prequantum bundle, (B,∇), we can use its geometric structure to construct a representation of the *entire* Poisson algebra of complex valued functions on M as an algebra of symmetric operators on $L^2(M,d\mu)$ (here dμ is the Liouville measure on M which is associated with ω). That is, the prequantum bundle determines a map such that .

Here we seem to have the beginnings of a quantum theory: we have constructed a Hilbert space, and an algebra of observables which corresponds to the *full* Poisson algebra of classical observables. This is, of course, too good to be true. The 'states' in our Hilbert space depend on 2n classical variables, whereas the states of the type of quantum theory that we know and love depend only on n classical variables.

(ii) Polarization. What we have to do, in effect, is to eliminate n variables from the arguments of our quantum states. In the case of the ordinary Schrödinger representation of quantum mechanics where M=T*Q, this corresponds to taking $L^2(Q)$ rather than $L^2(M)$ as our Hilbert space. In order to accomplish this, we must fix a *polarization*. This is a foliation of M by Lagrangian submanifolds.[106] The idea is that specifying a polarization is roughly equivalent to choosing a maximal set of Poisson-commuting classical variables. I.e., upto global topological problems, there is a one-to-one correspondence between polarizations of M and sets, $\{f_1,\ldots,f_n\}$, of classical observables with $\{f_i,f_j\}=0$. In the case where M=T*Q,

[105] This is a consequence of the fact that $H^1(M,\mathbf{R})$ is the abelian part of the fundamental group of M.
[106] N is a Lagrangian submanifold of M if $\omega|_N=0$ so that ω(v,w)=0 for all v,w∈T*N.

the cotangent bundle structure singles out the *vertical polarization* according to which M is foliated by manifolds of the form T^*_qQ for $q \in Q$.

We fix a polarization of M and then stipulate that our quantum Hilbert space is the subspace of the prequantum Hilbert space $L^2(M)$ which consists of square integrable functions on M which are covariantly constant along the leaves of the polarization (here we use the connection on the prequantum bundle). The states in this quantum Hilbert space now, in effect, depend on only n classical variables, since we have forbidden them to depend on $\{f_1,\ldots,f_n\}$—in the case of the Schrödinger representation, we simply kill the dependence on the momentum variables. We can now attempt to restrict our original representation of the classical Poisson algebra to this subspace. The result is that some, but not all, of the classical functions can be represented as operators on the quantum Hilbert space.

After having performed geometric quantization by carrying out prequantization and polarization, we are left with a Hilbert space and an algebra of observables which corresponds to a subalgebra of the Poisson algebra of classical observables. This takes us as far as Step 5 of the quantization schema of the previous section: we have not as yet represented all of the classical observables that we are interested in, nor have we implemented the dynamics. Indeed even if we are working with a cotangent bundle phase space, $M=T^*Q$, and select the canonical prequantum bundle and vertical polarization, we still find that the angular momentum and energy of our system are not yet represented by quantum observables. Completion of the quantization schema is not quite so elegant as the first two steps of geometric quantization—here things must be done on a case by case basis. Nonetheless, the program has been successfully completed for a large number of standard examples from ordinary quantum mechanics.

Geometric quantization isolates the input which is required to quantize a Hamiltonian system (M,ω,H): one needs to construct a prequantum bundle over M, and to select a polarization of M. Both of these steps have nontrivial physical content: the inequivalent prequantum bundles are parameterized by $H^1(M,U(1))$ and in general we find that quantizations based upon inequivalent prequantizations are inequivalent; furthermore, there are many polarizations for each M, and it is known that distinct polarizations can, but need not, lead to inequivalent quantum theories.[107]

Thus, generically, there are a multitude of physically inequivalent quantizations of any given Hamiltonian system (M,ω,H). How can this be reconciled with the Stone-von Neumann theorem? The theorem merely states that all of the irreducible representations of the Poisson algebra for a given set of canonical coordinates are unitarily equivalent. This leaves open the possibility that inequivalent quantizations of the same classical system can result from choosing two sets of canonical coordinates which satisfy distinct Poisson bracket relations. This possibility is realized whenever the classical phase space is non-simply connected.

But how seriously should we take this multitude of possible quantizations? It is tempting to adopt the following skeptical line: for some exotic classical systems, perhaps, each of the geometric quantizations will be interesting; but for run of the mill classical systems, of the sort found in textbooks, there is an obvious quantization which should be taken seriously; the rest are mere mathematical curiosities, of no particular relevance to either physics or philosophy of physics.

I will argue that this skeptical attitude is untenable. I will challenge the skeptics on their preferred ground. I will grant them that many interesting classical systems have phase spaces which are cotangent bundles, and therefore suggest natural choices for the prequantum bundle and polarization. But I will present an example of such a system where, despite the existence of a distinguished geometric quantization based upon the canonical prequantum bundle and the vertical polarization, there are nonetheless infinitely many other

[107] In ordinary quantum mechanics, the position and momentum representations are based upon distinct polarizations, but yield equivalent quantum theories; see §1.8 of Sniatycki 1980. In the worst case, however, choosing the ‘wrong’ polarization of a perfectly good classical phase space can lead to a trivial quantum Hilbert space; see Woodhouse 1980, pp. 140-41. I know of no detailed analysis of the conditions under which distinct polarizations lead to inequivalent quantum theories.

geometric quantizations which are physically relevant.[108] I will argue that this shows that the selection of a quantum theory has interesting interpretative repercussions.

To this end, let Q be $\mathbf{R}^3$ with an axis removed. Now suppose we are interested in the classical system which consists of a charged particle moving in the region external to a solenoid. We can model this system using a Hamiltonian system (M,ω_B,H) which is constructed as follows (see §5.11 of Woodhouse 1980; see also Horvathy 1980): we take $M=T^*Q$ as our phase space; we equip it with the symplectic structure $\omega_B=\omega+(e/2)\pi^*(*B)$, where ω is the canonical symplectic structure on T^*Q, e is the charge of the particle, π: $T^*Q\rightarrow Q$ is the projection map, and $*B$ is the Hodge star of the magnetic field, and H takes care of any non-electromagnetic forces acting on the particle.[109] So for a classical particle, it doesn't matter whether or not the solenoid is operational—either way, B=0 on Q so that $\omega_B=\omega$.

Now suppose we are interested in modeling the behavior of a quantum charged particle in the region external to the solenoid. That is, we want to quantize the particle while leaving the electromagnetic field classical. Then we are interested in quantizing our Hamiltonian system (T^*Q,ω,H). There are infinitely many geometric quantizations of this system, even if we stipulate that we are only interested in the vertical polarization. These quantizations are parameterized by $H^1(M,U(1))=U(1)$. So one can specify a geometric quantization by specifying a $z\in\mathbf{C}$ such that . Of course, one possibility stands out: we can chose $B=M\times\mathbf{C}$ with the canonical connection ∇. This corresponds to the identity element of $H^1(M,U(1))$.

Thus we have a perfect case for testing the skeptical claim. We have a physically interesting system which has infinitely many geometric quantizations. The skeptics insist that only one of these is relevant, the one based upon the distinguished canonical prequantum bundle. Are they correct?

---

108 These quantizations will be generated by using non-canonical prequantum bundles. I will assume throughout that we are only interested in vertical polarizations.
109 Recall that $*B$=,where the vector (B^x,B^y,B^z) is $g^{ab}B_a$.

No. We know from our discussion of the Bohm-Aharaonov effect in §6.8 that they cannot be. According to the distinguished quantization, an electron transported in a loop around the solenoid experiences no phase rotation, regardless of whether or not the solenoid is operational. Experiment reveals, however, that an electron which circumnavigates the solenoid will experience null phase rotation iff the solenoid is switched off. Thus the distinguished quantization is going to be an inadequate representation of our physical situation whenever the solenoid is working. In this case, we have to look to a geometric quantization of (M,ω,H) which is based on a non-canonical prequantum bundle. It turns out that we can view the parameter $z \in H^1(M,U(1))$ which determines the prequantum bundle as also determining the phase rotation of an electron which is transported around the solenoid.

Thus experiment tells us that we have to take seriously each of the infinitely many geometric quantizations of the classical system (M,ω,H). And this has interpretative consequences at the classical level. Historically, it was believed that the vector potential did not include any physical information beyond that which was contained in the electric and magnetic field vectors. This was perfectly plausible from the standpoint of the classical theory, since, as we saw above, the behavior of a classical particle can be predicted from knowledge of the electromagnetic tensor alone (even when the topology of space is non-trivial). Experiment reveals, however, that this viewpoint is not tenable when quantum particles are taken into account. The electric and magnetic field configurations fail to determine the behavior of a quantum particle. This should cause us to reconsider our understanding of the classical ontology, and to admit that there is physically relevant information encoded in the vector potential, above and beyond that captured by the electromagnetic field tensor. In fact, we can write

,

where $z \in U(1)$ is our quantization parameter which corresponds to the phase rotation of an electron transported around the solenoid, and A is the vector potential in the region external to the solenoid.

In this case, facts about quantization have interpretative repercussions. The following is a helpful way of thinking about how this works. Suppose that we are working in the pre-Aharonov-Bohm era, and are firmly committed to the interpretative view that the vector potential has no physical relevance. Our attempts to quantize the classical system (M,ω,H) will be guided by this view. If confronted with the multiplicity of possible geometric quantizations of the system, we will insist that at most one of them is of physical interests—most likely, of course, we will plump for the quantization based upon the canonical prequantum bundle. The discovery that each of the geometric quantizations is physically relevant should make us abandon this view. Thus we have a situation where an interpretative view about the classical theory mandates an approach to the technical problems of quantization; discovering that that particular approach to quantization is empirically untenable makes us abandon our views concerning classical ontology.

(3) Quantum field theory on a curved spacetime.[110] We fix a globally hyperbolic model of GR, (M,g). We are interested in quantizing a classical field, ϕ, with (M,g) as a fixed background. Thus we assume that we are dealing with a 'test field'—i.e., ϕ makes no contribution to the stress-energy tensor which determines the geometry of spacetime. This simplifies the task of quantization considerably, since it means that we do not have to worry about the interaction between the energy-momentum of the field and the geometry of spacetime when we quantize. Even so, the task is unmanageably difficult unless we impose the further restriction that ϕ be a *linear* field. Mathematically, this means that the phase space of the classical field theory is a vector space. Physically, this means that ϕ is not interacting with itself or any other physical field.

For definiteness, we will suppose that ϕ is a real Klein-Gordon field.[111] We now think of the Klein-Gordon field as representing a classical field on spacetime rather than as

[110] This example is long and complex. It can be omitted without loss of continuity. See Wald 1994 for an outstanding textbook treatment of this material.
[111] Everything in what follows goes through for other linear field theories. See §4.7 of Wald 1994.

being the wave function of a relativistic free particle.[112] We choose a foliation, parameterized by t, of (M,g) into Cauchy surfaces Σ_t diffeomorphic to some three manifold Σ, and an identification map $\theta_t:\Sigma_t\rightarrow\Sigma$ for each t. A point in the configuration space Q is picked out by specifying a field configuration at an instant in time, Σ_t. Thus $Q=\{\phi: \Sigma\rightarrow\mathbf{R}\}$. Our phase space is T*Q, with the symplectic structure, ω, induced by stipulating that the momentum conjugate to ϕ on Σ_t is

,

where q_{ab} is the Riemannian metric which g induces on Σ_t, and n^a is the unit normal to Σ_t in (M,g). We now choose a Hamiltonian H for (T*Q,ω) so that Hamilton's equations of motion give us the Klein-Gordon equation on (M,g):

.[113]

The Hamiltonian system (T*Q,ω,H) is our theory of the Klein-Gordon field on the background spacetime (M,g). The construction is independent of the choice of foliation, time function, and identification map in the sense that any such choice leads to a symplectic manifold (T*Q',ω') isomorphic to our (T*Q,ω), and to a Hamiltonian H' such that the fields on (M,g) which correspond to the dynamical trajectories of (T*Q',ω',H') are identical to those which correspond to the dynamical trajectories of (T*Q,ω,H).[114]

The Klein-Gordon equation is linear, so $\phi+\phi'$ is a solution if ϕ and ϕ' are. This means that Q and T*Q are vector spaces, so that we may identify T*Q with its tangent space, and regard the symplectic form ω as a bilinear map $\omega: T^*Q\times T^*Q\rightarrow\mathbf{R}$. This allows us to adopt the following approach to quantization, which works for any Hamiltonian system (M,ω,H), where (M,ω) is a symplectic vector space.

---

112 Thus we are engaging in field quantization rather than second quantization. See pp. 13-15 of Redhead 1990 or Chapter 4 of Teller 1995 a discussion of this distinction.

113 Compare with example 5.3. I forego writing down the Hamiltonian because it is a complicated expression involving the identification map.

114 The theory can also be formulated as a gauge theory which does not require a choice of foliation, time function, or identification map. See Kuchar 1981 or Hajicek and Isham 1996b.

As our input, we require a Hilbert space $\mathcal{H}$, which we will call the *one particle Hilbert space*. We now use $\mathcal{H}$ in constructing a far larger Hilbert space, the *Fock space*:

,

where is the direct sum, and is the symmetrized tensor product (see Appendix A.2 of Wald 1994; if we are dealing with a fermionic field, then we use the antisymmetrized tensor product in constructing our Fock space). Our Fock space is built out of an infinite number of copies of $\mathcal{H}$ in such a way that we can interpret certain states in $\mathcal{F}(\mathcal{H})$ as representing systems of n copies of the quantum system represented by $\mathcal{F}(\mathcal{H})$—this is why we call $\mathcal{H}$ the one particle Hilbert space (see Wald 1994 or Teller 1995 for the details). More precisely: given a basis $\{v_i\}$ for $\mathcal{H}$, we can write any vector in $\mathcal{F}(\mathcal{H})$ as $(n_1,n_2,n_3,\ldots)$, with ; this state corresponds to a superposition of n_1 one particle systems in state v_1, with n_2 one particle systems in state v_2, etc. In particular, there is a special state, =(0,0,0,…), called the *vacuum state* (is independent of the choice of basis for $\mathcal{H}$). It is important to keep in mind that, despite this nice feature, $\mathcal{F}(\mathcal{H})$ is an ordinary Hilbert space, and need not have anything to do with particles.

If we select our one particle Hilbert space judiciously, then it will be easy to find a representation of the classical Poisson algebra as an algebra of observables on our Fock space.[115] In order to describe the required construction, we must first consider the complexification of T*Q.[116] This is the complex vector space V=. We extend ω by linearity so that it is a bilinear form on V. The map (·,·): V×V→**C** given by (v,w)=iω(,w) is

---

[115] Because the phase space is a vector space, and hence is simply connected, there is no interesting ambiguity in selecting a subalgebra of the classical Poisson algebra. Equivalently: there is only one prequantum bundle over (M,ω).

[116] In fact, I cheat quite a bit in presenting this construction (see the discussion on pp. 40-42 of Wald 1994). The technique I present can only be made rigorous in the finite dimensional case. In the infinite dimensional case, one runs into difficulties caused by the fact that the space of classical solutions must be enlarged to include distributional solutions before quantization can proceed. This is, however, a strictly technical difficulty. I take the liberty of presenting this less-than-rigorous approach here, since it has the advantage of emphasizing the connection with geometric quantization (see the next footnote). This connection is present, but obscured, under the more rigorous approach (see also Chapter 7 of Woodhouse 1980).

almost an inner product on V—it falls short only because it isn't positive definite. We can, however, find subspaces of V upon which iω(,w) is an inner product. We now look for such a subspace of V' of V which satisfies the following three desiderata: (i) iω(,w) makes V' into a complex Hilbert space; (ii) , where is the complex conjugate of V' in V; (iii) iω(,w)=0 for v∈ V' and w∈ .[117] The inner product space $\mathcal{H}$=(V',iω(·,·)) is our one particle Hilbert space.

As promised, the Fock space $\mathcal{F}(\mathcal{H})$ caries an irreducible representation of a large subalgebra of the algebra of classical observables.[118] These are represented as creation and annihilation operators on $\mathcal{F}(\mathcal{H})$ (i.e., operators which change $(n_1,n_2,n_3,\ldots)$ to $(n_1,n_2,n_3,\ldots,n_j\pm 1,\ldots)$). Furthermore, there is a standard technique for implementing the dynamics of the theory. Thus it seems fair to say that a selection of a one particle Hilbert space completes most of the hard work required to quantize a linear classical theory.

The above sketch is very cursory. In particular, I have skimped on providing details about the sense in which a vector in the one particle Hilbert space represents a quantum particle, and the sense in which a vector in the Fock space represents a quantum field, but can also be viewed as a superposition of aggregates of systems represented by the one particle space. These omissions are unavoidable—to remedy them would be to write an introduction to the formalism and interpretation of quantum field theory. For present purposes, it is enough to make two points.

(i) The quantum theory that we have constructed has a familiar form: a Hilbert space of states, together with an algebra of self-adjoint observables, with dynamics implemented by a Schrödinger equation. We can go ahead and attempt to interpret this quantum theory as we would any other, without paying any attention to the fact that our Hilbert space happens to be a Fock space.

[117] Notice the similarity to polarization: we are, in effect, killing half of the classical variables by restricting attention to V'. Since there is no freedom in choosing a prequantum bundle, the choice of V' amounts to a choice of geometric quantization.

[118] Unfortunately, this algebra of quantum observables will not, in general, include the stress-energy of the field. This causes serious difficulties if one is interested in studying the interaction between the quantum field and the geometry of spacetime.

(ii) The particle interpretation is often illuminating. If we quantize Maxwell's theory in Minkowski spacetime, then the particles of the particle interpretation are photons. Nonetheless, it is not always wise to assume that because we have a Fock space, we have a theory about a potentially infinite collection of quantum particles. After all, our formalism applies to any linear theory—including the theory of a single classical harmonic oscillator. But when such a finite dimensional system is quantized we typically want to think of the result as describing a single quantum particle rather than a indefinite collection of such things.

Nonetheless, the status of particles has been the focus of philosophical discussion of quantum field theory (see, e.g., Teller 1995 and the papers in Brown and Harré 1988). Teller comes close to claiming that "quantum field theory" is a misnomer, on the grounds that such theories are in fact theories of collections of quantum particles (see especially pp. 93-94). This attitude is diametrically opposed to that of Wald, who maintains that:

> Quantum field theory is a quantum theory of *fields*, not particles. Although in appropriate circumstances a particle interpretation of the theory may be available, the notion of "particles" plays no fundamental role either in the formulation or interpretation of the theory. (Wald 1994, p. 2)

I will argue that this interpretative disagreement arises out of a divergence of reactions to the freedom available in selecting a quantization for a given classical field theory. Proponents of particles recognize that there are infinitely many quantizations of a given field theory in Minkowski spacetime, but claim that there is a single such quantum theory which is physically relevant. Wald. on the other hand, takes all of these quantizations equally seriously.

Since the fact that we are dealing with linear classical theories means that there is no freedom in selecting a classical Poisson algebra of observables, and since selection of a one particle Hilbert space leads directly to a representation of such an algebra as an algebra of operators on a Fock space as well as to a straightforward method for implementing the quantum dynamics, all of the ambiguity in quantization lies in the selection of a one particle Hilbert space. In finite dimensions it turns out that even this freedom is trivial—since the

Stone-von Neumann theorem guarantees that all irreducible representations of our algebra are unitarily isomorphic. In the infinite dimensional case this theorem does not apply, and we can have unitarily inequivalent representations of a given algebra. Indeed, different one particle Hilbert spaces can lead to such inequivalent quantum theories, so that our technique of quantization leads to an infinite number of inequivalent quantizations of any given classical field theory.

From a certain perspective, this freedom can appear to be quite benign. Let us suppose that we are interested in quantizing the Klein-Gordon field theory (T*Q,ω,H) in Minkowski spacetime. Then there is a natural choice for a one particle Hilbert space: the space of solutions to the Klein-Gordon equation which oscillate with positive frequency relative to the time associated with any inertial frame (see example 1).[119] Similar procedures give us an unambiguous procedure for quantizing any linear field theory set in Minkowski space. We can then go on to utilize the particle interpretation of the Fock space, and to speak of the quantization of Maxwell's theory as being the theory of free photons, etc.

This is the point of view adopted in most introductory textbook treatments of quantum field theory: (i) quantum field theory involves the quantization of a classical field living in Minkowski spacetime; (ii) here it is possible to formulate an unambiguous quantization procedure; (iii) so it is natural to base an interpretation of the theory upon a particle interpretation of the resulting Fock space. The bulk of the philosophical literature on quantum field theory can be read as an attempt to work out the details of this approach, or to criticize (iii) without questioning (i) and (ii). It is not, of course, that philosophers are unaware that (i) is false in the sense that there is much interesting work on quantum field theory on curved spacetimes, nor that they are unaware that the status of (ii) is, as we will see, subtle. Rather it is that they have chosen to work on a carefully circumscribed theory—essentially that of the textbook presentations. In doing so, they are both secure in the knowledge that even this thin notion of quantum field theory presents more than enough

---

[119] Equivalently: there is a natural polarization of T*Q; see §7.10 of Woodhouse 1980.

interpretative difficulties, and hopeful that their work on this theory will carry over to a more general setting. The issue between Teller and Wald is whether or not this strategy is paying off.

In order to mediate this dispute, we will need some facts about quantum field theory on a curved spacetime.

Let's begin by considering the theory of the Klein-Gordon field on a generic general relativistic spacetime (M,g) with no symmetries. In this case, the geometry of spacetime does not provide enough structure to construct a distinguished one particle Hilbert space—one such space is as good as another. Given two one particle Hilbert spaces, $\mathcal{H}_1$ and $\mathcal{H}_2$, we can look at the quantum theories to which they lead. These will be set in the Fock spaces $\mathcal{F}(\mathcal{H}_1)$ and $\mathcal{F}(\mathcal{H}_1)$. There is a canonical technique for constructing a correspondence, U, between the two quantum theories (see §4.4 of Wald 1994). If the quantum theories are unitarily equivalent, then U: $\mathcal{F}(\mathcal{H}_1)\rightarrow\mathcal{F}(\mathcal{H}_2)$ is an isomorphism between the theories. If they fail to be unitarily equivalent, then U maps vectors in $\mathcal{F}(\mathcal{H}_1)$ to algebraic states (see §2 above). These are not vectors in $\mathcal{F}(\mathcal{H}_2)$, or even density matrices, so they cannot be regarded as true states of the second quantum theory. Nonetheless, they can be associated with outcomes of measurements of the observables of the theory, and hence are of physical interest. In particular, if the theories are not equivalent, then not only do we have that U($_1$)≠$_2$, we also have that U($_1$) is not a true state of the second quantum theory. This is ordinarily understood as saying that the two theories do not even agree about what the vacuum looks like.

Now, how much structure does (M,g) have to have in order to single out a natural choice of a one particle Hilbert space? The best answer appears to be that (M,g) must admit a timelike Killing field; i.e., it must admit a one parameter family of timelike isometries. When such a Killing field exists, we can use the associated isometries to impose a preferred notion of time on the spacetime, and this in turn allows us to single out a one particle Hilbert space which (roughly) corresponds to the set of solutions to the classical field equations which propagate with positive frequency relative to this Killing time. Thus, if we restrict our attention to the class of spacetimes which admit a single Killing field, then we have a way of picking out a unique one particle Hilbert space, so that there is a preferred quantization for a field theory on such a spacetime.

If, however, our spacetime admits more than one Killing field, then we are still in our original situation—our geometry does not pick out a unique quantization, and distinct quantizations will typically be unitarily inequivalent, and thus involve distinct notions of 'particle' and 'vacuum.' Something like this happens with Minkowski spacetime. If we restrict our attention to the Rindler wedge W, given in inertial coordinates by $W=\{(x_0,x_1,x_2,x_3)\in \mathbf{R}^4: \}$, then we find that there are two Killing fields: the one determined by the family of time translations relative to any inertial frame, and the one determined by a one parameter family of boosts in the x_1 direction (there are, of course, infinitely many such families). We let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the one particle Hilbert spaces associated with the first and second Killing fields, respectively. These one particle Hilbert spaces lead to inequivalent quantum field theories. The first, of course, leads to the standard quantum field theory on Minkowski spacetime. The second leads to a nonstandard theory with a nonstandard notion of vacuum and particle. This is usually interpreted as being the quantum field theory which uniformly accelerating observer in Minkowski spacetime would write down. If we now look at the algebraic state, $U(_1)$, of the nonstandard theory which corresponds to the standard vacuum, we find that it is a thermal state. That is, the standard vacuum feels like a heat bath to an observer in the nonstandard theory.[120] This is known as the *Unruh effect*.

I believe that Wald's point of view is something like the following.[121] The existence of such radically inequivalent quantizations of classical field theories on arbitrary

[120] This provides a good example of the kind of true quantum chaos that is rare in quantum mechanics but common in quantum field theory; see Emch, Narnhofer, Thirring, and Sewell 1994.
[121] In addition to being an outstanding textbook, Wald 1994 also contains some very good philosophy of physics. Since, however, Wald's interpretative arguments are spread throughout the entire work and tend to be intertwined with technical discussions, I will not attempt to provide textual support for my exposition of his view.

spacetimes leaves us with a dilemma. We either maintain that there is one quantization which is *the* correct quantization of the classical theory; or we attempt to treat all of the competing quantizations equally seriously. Wald shows no interest in the former option—and, given the fact that there is in general no classical structure which can single out such a preferred quantization, it seems that there is little reason to do so. He pursues the latter option by advocating a broadly algebraic approach: the physical content of quantum field theory on a curved background is the intersection of the contents of each of the unitarily inequivalent quantizations. In particular, since the quantizations disagree about what a particle is, and what the vacuum state is, we should abandon these notions in interpreting the theory. We can still, of course, talk about $_1$: it is a perfectly legitimate state in the quantum theory in $\mathcal{F}(\mathcal{H}_1)$, and it corresponds to a legitimate algebraic state, $U(_1)$, according to the second quantum theory.

When it comes to spacetimes which do happen to admit a preferred quantization, Wald sticks to his guns and continues to push an algebraic approach. That is, he advocates a univocal interpretation of quantum field theory on a fixed background spacetime: since there is no preferred notion of particle available for most spacetimes, we should not base our interpretation on such a notion *even when one is available*. He bolsters this view by pushing an analogy between the choice of a particular quantization of a classical field, and the choice of a coordinate system in GR (see, e.g., pp. 2, 60). In both cases, we are meant to conclude, a specific choice of structure may be helpful for a given calculation, but when we turn to interpretation we should only take seriously those structures that are independent of such a choice. Thus, just as we do not draw any conclusions about ontology from the fact that some spacetimes admit preferred sets of coordinates, so we should not be overly impressed by the fact that some spacetimes permit us to pick out a distinguished concept of particle.

This view of Wald's stands in opposition to two well-established approaches to interpreting quantum field theories. The first I will call the *conservative* approach. This is perhaps endorsed by the majority of philosophers (and physicists) who work on quantum field theory in Minkowski spacetime alone. It is widely held that one should take what one can get in this case: since there is a preferred notion of particle in Minkowski spacetime, one should employ it—notwithstanding the difficulties which the notion of particle encounters in the more general setting. This conservative strategy amounts to an insistence that it is legitimate to take quantum field theory in Minkowski spacetime as constituting a legitimate theory in its own right. As such, its problems of interpretation may be conceived to be free-standing: the fact that there are problems with the notion of particle in some other theory (quantum field theory on a curved spacetime), no more impugns the concept of particle in Minkowski spacetime, than the fact that results from high energy physics suggest that spacetime has a discrete structure can force us to abandon attempts to interpret GR in terms of a continuous physical spacetime (see §1.4).

This is a consistent alternative to Wald's position. In §2.12, I claimed that Machian relationalism about the space of nonrelativistic physics is triumphant if the assigned task is to interpret an observationally adequate fragment of Newtonian physics. This was not meant to be an exotic counterfactual. One might well be interested in interpreting only as much of the Newtonian theory as is needed—especially when taking the remainder into account is enough to disqualify a very elegant interpretative strategy. The same move is certainly possible with quantum field theory, although perhaps not as well motivated by philosophical tradition. From the perspective of this conservative strategy, Wald's rhetorical analogy between the choice of a set of coordinates in GR and the choice of a quantum field theory will appear to be sheer bluster—one need not feel any more ashamed of relying upon the flat structure of Minkowski spacetime to single out a preferred quantization than one need feel ashamed of taking the existence of a set of preferred inertial frames to be an important datum for any interpretation of special relativity.

This conservative strategy is not, however, an attractive option in the case at hand. The move requires that we insist that the quantum field theory based upon the standard time of Minkowski space is the only interesting quantum field theory—otherwise we would be unable to find a privileged notion of particle. But this means that we must go on to claim that the Unruh effect is a mathematical artifact, of no physical interest. For recall that the effect consists of an accelerating observer seeing the standard vacuum as a heat bath. This result is *not* derivable within the standard theory. That is, one cannot understand the Unruh effect along the lines of, say, time dilation—it does not follow from the standard theory of an inertial observer that an accelerating observer will undergo this effect.[122] Therefore, the Unruh effect can only be derived by taking the quantum field theory of the accelerated observer seriously. Thus the Unruh effect has no place in the approach to quantum field theory which we are considering. Now, I claim that this, far from being a beneficial side effect of this interpretative approach, is actually its downfall. The reason is straightforward. Many people *do* take the Unruh effect seriously, and with good reason: it is closely related to the Hawking effect, which, in turn, is widely considered to be the deepest available result concerning the interaction between the quantum and the gravitational (for the first claim, see, e.g., Connes and Rovelli 1994, Haag 1992, and Wald 1994; for the second, see, e.g., Ashtekar 1995, Isham 1991, and Jacobson 1991).

I believe that this is a sufficient reason to reject the conservative strategy. There is, however, a second option for someone who wishes to interpret quantum field theory in terms of particles. This has been given its canonical formulation in Teller 1995. One can stipulate that the particles of every quantization of a given classical field theory are to be taken equally seriously. In the Minkowski case, this means living with many kinds of particles: the standard Poincaré covariant particles of the quantization based upon inertial time, as well as the various kinds of *Rindler particles* which arise in the quantum field theories associated with families of accelerating observers. One need not go on to extend this interpretative strategy to quantum field theories in curved spacetimes. Teller, however, does so, and ends up advocating the existence of a distinct type of particle for every type of

[122] If this result could be derived within the standard theory, then the theory of the standard and accelerated observers would be unitarily equivalent. But we know that this is not the case. See Unruh and Wald 1984 for an analysis of the event of the measurement of a Rindler quanta within the standard quantum field theory.

observer in each general relativistic spacetime (see pp. 110-13 of Teller 1995). Nonetheless, he continues to maintain that particles are the primary resource for interpreting quantum field theories. He recognizes that this is a less than ideal interpretative strategy:

> One may sensibly charge that by this time we have wandered off so far from the original concept of particles that the concept has been given up entirely. We reject exact trajectories and … primitive thisness. … Now we find in addition that there are various "kinds" of quanta, each relative to the state of motion of an observer. All inertial observers in flat space-time agree, but not accelerating observers or observers following accelerated trajectories or geodesics in curved spacetime. …
>
> If this counts as completely giving up the particle concept, so be it. The terminology is not important. The important point is that we have traced the lineage of the particle concept, noting similarities and differences along the way, and we see more clearly how the conceptual parts fit together in the new theoretical framework. After that a dispute about the appropriateness of terminology carries little interest.[123]

But this is to miss the point of a position like Wald's. Teller grants Wald that the notion of particle is observer-relative, even in Minkowski spacetime. He nevertheless maintains that this notion is fundamental to the enterprise of interpreting quantum field theory. He expects Wald to retort that this notion no longer deserves to be called a notion of particle, since whatever particles are they are not observer-relative. Teller sees that this sort of objection can be defused if he is willing to concede the use of his term of art. But Teller has missed the real objection: it is not that Teller's 'particles' are not worthy of the name since they are frame dependent, it is that they can not be fundamental to the interpretative enterprise since they are frame dependent. This is where Wald's rhetorical comparison between the choice of a coordinate system and the choice of a quantum field theory is brought to bear. His position is that we should no more base an interpretation of quantum field theory upon a notion which is relative to a choice of one particle Hilbert space than we should base an interpretation of GR upon a notion, like the distinction between inertial and gravitational force, which is relative to a choice of coordinates in GR (see the discussion of harmonic coordinates in §6.7).

I find this line which I attribute to Wald quite convincing. And so I have the following view of the relation between quantization and interpretation in the case of linear field theories. Historically, the quantum theory of linear fields was first developed within the

---

[123] Teller 1995, p. 112.

special context of Minkowski spacetime. Within this context, there is a distinguished quantum field theory, which it was natural to assume was the only physically interesting quantization of the classical field theory. Proceeding under this assumption, people worked very hard on articulating a particle interpretation of quantum field theory. It turns out, however, that there are infinitely many inequivalent quantizations of any classical field theory on a fixed general relativistic background spacetime. Each of these quantizations employs a distinct notion of particle. This is true even in Minkowski spacetime.

At this point, there are a number of possible interpretative strategies. One can attempt to give a univocal interpretation of all of quantum field theory on curved spacetime, including the Minkowski case. Here there are three options: (i) follow Wald in attempting to give an interpretation which is independent of the choice of one particle Hilbert space—this means eschewing the notion of particle altogether; (ii) maintain that there is single correct choice of one particle Hilbert space for each classical field theory and each spacetime—which means that there is a distinguished notion of particle for each quantum theory of fields; (iii) follow Teller in attempting to hold on to the notion of particle, while treating each one particle Hilbert space equally. No one seems to take the second option seriously in the absence of any classical structure sufficient to single out a choice of quantum theory. And, as argued above, Teller's approach is unattractive. Thus Wald's approach is the most promising univocal interpretative strategy.

One need not, however, adopt a univocal strategy. One can take the conservative route, and employ one strategy in handling quantum field theories set in Minkowski spacetime, and another for theories set in curved spacetimes. Thus, one could employ Wald's strategy in the general case, but insist that there is one true quantization of a field theory set in Minkowski space. This allows one to maintain that there is distinguished notion of particle iff spacetime is flat (or, more generally, stationary). This is, in effect, to insist that quantum field theory in Minkowski spacetime, although a subtheory of quantum field theory on a general relativistic spacetime, is nonetheless a theory in its own right, and thus supports interpretative options which are unavailable in the larger theory.

Thus it looks as though we have a standoff between and those who, like Wald, advocate a univocal interpretation of quantum field theory which is appropriate for *any* spacetime background, and those who insist on the autonomy of quantum field theory in Minkowski spacetime. This deadlock is broken by the intrusion of concerns external to quantum field theory in Minkowski spacetime—it turns out that the approach which treats the subtheory as a theory in its own right for the purposes of interpretation leads one to disregard the Unruh effect. But, for reasons having to do with the general case of quantum field theory on a curved spacetime and with quantum gravity, many physicists are unwilling to take this course.

Thus we have a case where, taken on its own terms, a theory is able to resolve the ambiguity involved in quantization, despite the existence of infinitely many inequivalent quantizations. But this resolution depends on keeping the project of interpreting this theory isolated from the interpretative problems of a larger theory of which it is a subtheory. For reasons external to the subtheory itself, this turns out to be undesirable, so that the ambiguities associated with the complete infinite family of inequivalent quantum field theories must be faced.

4. The Standard Approach to Quantum Gravity, and the Problem of Time.

The last two examples of the previous section provide us with instances where the ambiguity associated with quantization led to an interesting interpretative interaction between two theories. In example 2, the interaction was between the classical theory and its quantization. In example 3, it was between two quantum theories, one of which was a special case of the other. The moral is that one should not be surprised to find difficult interpretative questions intertwined with the technical details of quantization. With this in mind, it is time to move onto the central example of this Chapter.

The examples of the previous section were in closed form: we had a classification of the quantizations available, and a feel for which interpretative moves are associated with which approaches to quantization. In the balance of this Chapter, I want to focus on the interpretative interplay between classical and quantum gravity. Here the story is open-ended: we do not have a single quantization of full GR; and the work of analyzing the interpretative ramifications of this process of quantization has not yet really begun. For this reason, I will have to resort to thought experiments when it comes time to explicate the interpretative interplay between classical and quantum gravity—if *this* approach to quantization were to win out, how would this fact reflect on *that* interpretation of GR?

We do have several sketches of quantizations of GR. Even when, as I do here, we restrict attention to canonical quantization, there is a large number of proposals for how to apply the general quantization schema of §2 to the theory. It is impossible to survey all of these approaches here (see Isham 1991, Isham 1993, and Kuchar 1992). Instead, I will sketch what I call the 'standard approach.' This approach is an attempt to apply the quantization schema of §2 to the formalism of §7.2 in the most straightforward possible manner. This approach has not met with a great deal of success. It runs into a number of very difficult technical problems. There are also a number of outstanding conceptual problems, which are our main interest here. Nonetheless, it is a useful starting point. Historically, the present research program of canonical quantum gravity grew out the original formulations of the standard approach in the work of Arnowitt, Deser, and Misner 1962 and Dirac 1964. And it remains true today that most approaches to canonical quantum gravity can be seen as variations on this theme. Some of these variations are motivated by technical concerns—and these we need not bother about here. But many are motivated by a hunch about how the conceptual problems facing the standard approach can be overcome. These are exactly the sort of variations that should interest us, since they are often intimately related to interpretative moves at the classical or quantum level. Thus the standard approach supplies a helpful frame of reference for discussion of present research. In terms of the seven steps of the quantization schema of §2, the standard approach proceeds as follows:

Step (1). Choose a classical formulation. As our classical theory, we take the full phase space formulation of GR outlined in §7.2. That is, we conceive of GR as a constraint surface $(N,\omega|_N,0)$ in the symplectic manifold (T^*Q,ω). Here Q is Riem Σ, the space of Riemannian metrics on some fixed three manifold, Σ, and N is determined by the scalar and vector constraints, and . We will call this *geometrodynamics.*

This decision is substantive in a number of ways. First of all, it means that we have decided *not* to fix the gauge, or reduce the phase space by eliminating any of the gauge freedom associated with the constraints.

Notice, furthermore, that the restrictions noted in §7.1 are still in force—we are restricting attention to the set of globally hyperbolic solutions to the vacuum EFE which are spatially compact. The restriction to globally hyperbolic solutions appears to be more or less necessary, at least for the time being—little progress has been made on the canonical formalism for non-globally hyperbolic GR. The restriction to spatially compact solutions is a technical convenience—some progress has been made on quantum gravity in the asymptotically flat case, but it is more convenient to discuss the simpler compact case. The restriction to vacuum solutions is substantive in two ways. First of all, the inclusion of phenomenological dust or fluids can lead to simplifications of the classical theory which are formally equivalent gauge fixation. In brief, this occurs because the worldlines and proper times of the dust particles, or whatever, give us a preferred set of coordinates on spacetime (see Kuchar 1993b, Rovelli 1991c, Rovelli 1991e, and Unruh and Wald 1989 for explorations of this strategy). Second, in some quarters, there is a feeling that it is impossible to find a consistent theory of quantum gravity which is not also a unified field theory which takes the other fundamental interactions into account (this sentiment is allegedly widespread among particle physicists; see §3 of Isham 1991 for a discussion of this sort of approach). Thus the decision to concentrate on models of GR which are matter-free is a substantive choice (see p. 280 of Ashtekar 1995 for a defense of this choice).

Finally, our decision to employ the formulation of §7.2 is a decision not to employ the Ashtekar variables (see Ashtekar 1991 or Ashtekar 1995). These variables are based upon a formulation of GR as a gauge theory set in phase space based upon a configuration space whose points are connections on a bundle over Σ rather than metrics on Σ. This formulation of GR as a theory of *connection dynamics* rather than geometrodynamics is perspicacious in a number of ways. Most importantly: (i) it casts GR into a form which in which the similarities and differences between GR and Yang-Mills theories are very apparent, so that techniques from other quantum field theories can be more easily adapted for the gravitational case; (ii) it casts the constraints into a much simpler form. For these reasons, recent years have seen a great deal of progress on the technical problems of quantum gravity in terms of the Ashtekar variables (see Ashtekar 1995). However, for the most part, it is true that quantum gravity based upon connection dynamics faces exactly the same conceptual problems as quantum gravity based upon geometrodynamics (see especially Kuchar 1993a).[124] So, for present purposes, it is safe to restrict our attention to the geometrodynamical program.

Step (2). Choose fundamental variables. There is a natural choice of fundamental variables within geometrodynamics: the tensors p and q on Σ which specify points in the phase space of GR. The algebra $\mathcal{A}$ of the fundamental variables is constructed by enforcing the Poisson bracket relations between p and q, as well as whatever additional algebraic relations are required to take account of the global structure of the phase space. The Poisson bracket relations are just:

[124] An interesting, and technically crucial, difference between the geometrodynamical and connection dynamic approaches is that the latter (but not the former) naturally leads one to think of classical and quantum field theories in terms of values on loops and knots embedded in spacetime rather than in terms of values at the points of space. This difference promises to be a rich source of interpretative puzzles and prizes.

$\{q_{ab}(x),q_{ab}(x')\}=0$, $\{p_{ab}(x),p_{ab}(x')\}=0$, and $\{q_{ab}(x),p_{ab}(x')\}=\delta^c{}_{(a}\delta_{b)}{}^d\delta(x,x')$,

where $x,x'\in\Sigma$. The remaining relations are more subtle—one will, for instance, want to impose some conditions to ensure that q is positive definite, but this is not entirely straightforward (see §4.4.2 of Isham 1991).

Step (3) represent the algebra $\mathcal{A}$. The idea is to mimic the ordinary Schrödinger representation of quantum mechanics. To this end, we take our vector space, V, to be the set of smooth complex-valued functions on Q=Riem Σ. We now want to represent the position variables as multiplication operators on this space, and the momentum variables as a derivative operators on this space. To do so, we write down the equations:[125]

.

Here our states, $\psi[q]$, are functionals on the space Riem Σ of metrics on Σ. When hit by a position operator, , a state is multiplied by the corresponding classical potion variable, q. When hit by a momentum operator, , the state is transformed by a partial derivative in the position variable conjugate to the classical momentum variable p.

This is a powerful heuristic scheme. It is important to note, however, that no one knows how to actually carry out this construction. One very serious problem is that for technical reasons one expects the states of quantum gravity to be functionals on the space of distributional metrics rather than on Riem Σ itself; but it isn't clear how this can be achieved, since Riem Σ isn't a vector space (see §4.4.5 of Isham 1991). Nonetheless, this approach to representing the algebra of geometrodynamical variables has played an important role in discussions of quantum gravity. One says that it is useful at the *formal* level—Although one does not know how to make it rigorous, one *does* know how to manipulate the symbols.

---

[125] The quantity on the right hand side of the second equation cannot be any ordinary derivative—it must be a functional derivative; see Appendix E.1 of Wald 1984 or §II.A.1 of Choquet-Bruhat, DeWitt-Morrette, and Dillard-Bleick 1982.

Step (4) Impose constraints on the set of states. We need to find operators $_a$ and , which correspond to the classical constraints

and .

We then impose the quantum constraints $_a\psi[q]=0$ and $\psi[q]=0$ in order to find the space of physical states, V_{phys}.

At the formal level, the vector constraint is no problem. We want to put 'hats' on the symbols appearing in the classical constraint. Because the form of the classical constraint is so simple, there is only one way to do this: . Furthermore, still working merely formally, it is easy to interpret the meaning of this constraint: it requires that the states be diffeomorphism invariant functionals on Riem Σ. That is, $_a\psi[q]=0$ iff $\psi[q]$ is such that if d: $\Sigma \rightarrow \Sigma$ is any diffeomorphism, then $\psi[q]=\psi[d^*q]$. The upshot of this approach is that states become functionals on Riem/Diff Σ. (See §4.4.5 of Isham 1991 for a nice discussion of the quantum version of the vector constraint and of some of the obstacles standing in the way of rigorizing this discussion).

Even at the formal level, there is a severe factor ordering problem with the scalar constraint. That is, because our classical expression is algebraically nontrivial, there are multiple ways to put hats on the classical terms. If, somewhat arbitrarily, we choose to place the 's to the right of the 's, then the quantum scalar constraint $\psi[q]=0$ becomes:

,

where is the *DeWitt supermetric*, and R(q) is the scalar curvature of q. In this form, the quantum scalar constraint is known as the *Wheeler-DeWitt equation*. There are serious technical obstacles standing in the way of giving a rigorous treatment of this equation (see §4.4.6 of Isham 1991). In addition, unlike in the case of the quantum vector constraint, the physical content of the Wheeler-DeWitt equation is far from evident. This is, in some ways, the root of all of the conceptual problems of quantum gravity. See step (7) below.

Step (5) Inner product. Let's suppose that we somehow manage to overcome all of the technical problems standing in the way of carrying out the last two steps. Then we will have succeeded in constructing the vector space V_{phys} of physical states of quantum gravity. Now we look for an inner product which makes this vector into a Hilbert space. There is considerable physical content in this choice—ultimately, presumably, we will want to use our quantum theory to calculate probabilities, and these probabilities will depend as much on our choice of inner product as on the physical states involved. Not surprisingly, then, the selection of an inner product for quantum gravity is an extremely difficult task. This is one of the most imposing challenges facing the standard approach.

In meeting this challenge, we have one guiding principle, which is taken over from more familiar quantum theories: the inner product should make the observables self-adjoint.[126] This is of only limited value however: we will see that what observables one includes in the quantum theory depends upon one's approach to the problem of time. Thus, on one influential view, to be in a position to choose an inner product is to have solved the majority of the outstanding conceptual and technical problems of canonical quantum gravity (see Kuchar 1992 and Kuchar 1993a for this sort of line).

Step (6). Observables. At this stage, one wants to select a family of self-adjoint operators on the Hilbert space which correspond to measurable (or beable) quantities. Conventional wisdom has it that these operators should commute with the quantum constraints, so that the observables of the quantum theory are gauge invariant. Thus, they should correspond to gauge invariant classical observables.

Within the geometrodynamical program, this is a very tall order. The fundamental variables of this approach, p and q, are not gauge invariant. So one must first look for some

[126] Here the connection dynamics framework enjoys a considerable advantage over the geometrodynamical approach. The fundamental variables of connection dynamics are complex valued, so that their algebra includes an operation corresponding to complex conjugation. There is some hope that stipulating that the inner product on the space of physical states must respect this operation will single out a unique inner product. See Rendall 1994.

gauge invariant functions on the classical phase space which correspond to measurable (or beable) quantities. One must then find an appropriate representation of these quantities as operators on the Hilbert space. It turns out that even the first step is prohibitively difficult. No one knows how to write down an explicit expression for a gauge invariant observable on the classical phase space. Something even worse is true if we impose the reasonable-seeming condition that the observables must be local (i.e. that they must be expressible as integrals over Σ of p and q and their derivatives): there are no nontrivial local observables (see Torre 1993).

Step (7) Dynamics. At this stage we are forced to grapple with the fact that the general covariance of GR implies that the theory is time reparameterization invariant, which in turn implies that the quantum theory is prima facie timeless. This is the problem of time. According to the standard approach, quantum gravity consists of a Hilbert space of states together with an algebra of observables represented as operators on this space. But there is no dynamical equation, so that there is no change and no time in the theory (this follows from the fact that the Hamiltonian of GR is a sum of the constraints—see §2 above). This makes it extremely difficult to see how we are supposed to use the theory to model our world, in which time plays so conspicuous a role.

This means that we cannot employ quantum gravity as we do other quantum theories—we cannot fix a present state and make predictions about future states by imposing some sort of dynamics upon our system. On a straightforward reading of the standard approach to quantum gravity, there is no dynamical evolution. Furthermore, it is not at all clear how we can make sense of the unchanging states of the system. Naively, we expect that, since a state is of the form $\psi[q]$, the inner product on $\mathcal{H}$ should allow us to calculate the probability of finding some given three geometry (Σ, q) if we make an appropriate measurement. But how can we make sense of this probability? We believe that we live in a universe whose geometry changes radically over time, so that this time invariant probability of finding a given spatial geometry does not seem to be relevant to any empirical determination that we could make of the geometry of space.

The problem of time is the central conceptual problem of quantum gravity.[127] The standard approach to quantum gravity results in a theory which is very difficult to recognize as a description of our world unless one can find time hidden somewhere within the formalism. Most physicists are confident that this can be done, although it may require a departure from the standard approach. This confidence stems from the following counting argument (see Barbour 1994a and Barbour 1994b for a dissenting opinion). Consider the configuration space of GR. Points of this space are just Riemannian metrics, q_{ab}, on the three manifold Σ. Each such q_{ab} has six independent components at each point $x \in \Sigma$. But this does not mean that GR has six degrees of freedom per point of Σ, since the q_{ab} are subject to the constraints. Three degrees of freedom are removed by imposing the vector constraint (since this is equivalent to factoring out by the group of diffeomorphisms on the three manifold Σ). This means that the points of superspace represent three degrees of freedom per point of Σ. Now, the weak field limit of GR is the theory of a spin-2 field with two degrees of freedom per point of space (see §4.4b of Wald 1984). So, one concludes that the gravitational field of GR itself has two degrees of freedom per point of space. Therefore: time corresponds to one of the three degrees of freedom remaining after we have factored out the gauge freedom corresponding to the vector constraint. This, of course, is what we should expect from our intuitive picture that the scalar constraint generates the temporal evolution of spatial geometry. Thus, if we look at a quantum theory based upon the superspace formulation of GR (where we eliminate the gauge freedom associated with the vector constraint at the classical level), then our wave functions will depend upon three degrees of freedom per point of space, and we know that only two of these degrees of freedom correspond to the freedom to specify an instantaneous state of the gravitational

[127] As we will see, it is closely related to other conceptual problems, such as the nature of the observables of the theory.

field. Therefore, we should think of the states ψ[q] as depending on *time* as well as the state of the field.

The trick is to make this dependence explicit. Here our model is provided by the Newtonian and relativistic particle (see example 1 above). There, as in the case of GR, the constraints govern the dynamics. In the case of the Newtonian particle, the constraint $H'p_t+H=0$ is linear in the momentum conjugate to time, so that the quantum constraint ψ=0 is just the Schrödinger equation,

.

This is an example where time is hidden among the canonical variables, but is extremely easy to recover. In the case of the relativistic particle, the constraint/Hamiltonian is =0. Since this is not linear in the momenta, that quantum constraint is *not* a Schrödinger equation. Rather, it is the Klein-Gordon equation, . In this case, however, time is again hidden: we can transform the Klein-Gordon into a Schrödinger equation by exploiting the temporal symmetries of the Minkowski metric. In this case, it requires a good deal of work to recover time, but it can be done. This is the sort of conjuring trick that we would like to accomplish with quantum gravity: we would like to find time among the fundamental variables of the geometrodynamics, so that we could rewrite the scalar constraint as $h=p_t+H$, and replace the Wheeler-DeWitt equation with a Schrödinger equation:

,

where φ and π are fields on Σ which correspond to the two physical degrees of freedom (per point of space) of the gravitational field.[128]

This is a very tall order, however. If it is true that time is hidden among the canonical variables, then it is deep undercover. It is not, for instance, possible to apply the techniques which were developed to deal with the Klein-Gordon field. Recall that those techniques

[128] I simplify greatly here. Presumably, we will not want to privilege a single time parameter for quantum gravity, since we will still want to recognize the admissibility of arbitrary slicings in GR. Thus we will actually be looking for a many fingered Schrödinger equation. See §3.4.3 of Isham 1993, or §4 of Kuchar 1992.

could be generalized to the general relativistic case only when the background spacetime against which the field propagated was stationary. In the case of quantum gravity, the analogous requirement is that the DeWitt supermetric on superspace should admit a Killing field—which it does not (see Kuchar 1991).

5. Two Proposed Solutions to the Problem of Time.

There is, at present, no generally accepted solution to the problem of time. Indeed, I do not even know of anyone who claims to posses a fully worked out solution. On even the most optimistic evaluation of the situation, it will be impossible to apply or interpret a theory of quantum gravity until the problem of time is solved (see, e.g., Ashtekar 1995 or Carlip 1990 for this sort of attitude). According to the most pessimistic estimate, it will not even be possible to formulate a theory of quantum gravity until the problem of time is solved, since the definition of the observables and innerproduct of the theory are dependent upon one's approach to time (see, e.g., Kuchar 1992 and Kuchar 1993a).

It is not possible to offer an overview of even the most important proposed solutions to the problem of time in this Chapter—that project lies in the realm of the 100 page survey article (see Isham 1991, Isham 1993, and Kuchar 1992). That would not, in any case, serve my purposes particularly well. My project is to demonstrate that there is an interpretative interaction between classical and quantum gravity. My best hope for accomplishing this task is to focus on a couple of perspicacious examples, rather than to bury the reader under an avalanche of proposed solutions to the problem of time.

To this end, I will discuss two proposed solutions in this section, chosen because they present a particularly clear view of the interdependence between the interpretative problems of GR and approaches to the problem of time in quantum gravity. I am confident that they are atypical, not in exhibiting the classical-quantum interpretative interplay, but only in exhibiting it so clearly.

Examples

(4) Time as mean curvature (see Carlip 1990 or Beig 1994). We begin with what is perhaps the clearest instance of an interpretative interaction between classical and quantum gravity. As usual, we fix a three manifold Σ and view GR as a constraint surface $(N,\omega,0)$ in T^*Q, where $Q=\text{Riem } \Sigma$. Given $(p,q)\in N$, we can construct the *mean curvature*, $\tau=$. We will say that (q,p) is a *slice of constant mean curvature*, or a *CMC slice*, if τ is constant on Σ. We say that a spacetime (M,g) is CMC sliceable if it can be foliated by CMC surfaces. The set of CMC sliceable spacetimes is an open set, of unknown extent, in the space of models of GR.[129] The wonderful thing about CMC sliceable spacetimes is that τ varies monotonically within a foliation by CMC slices. This suggests that we can use the mean curvature τ as a time parameter.

This is indeed possible. We proceed as follows.

(i) We eliminate the gauge freedom associated with the vector constraint by adopting the superspace formulation of GR, $(N,\omega,0)$.

(ii) We restrict our attention to the set $N'\subset N$ of CMC slices. This combines two moves. The first is to discard all models of GR which are not CMC sliceable. In what follows we disregard any physical content this move may have.[130] The second step is a partial gauge fixation: one excludes all of the non-CMC slices which are Cauchy surfaces of the CMC sliceable spacetimes. This amounts to a commitment to describe the models of GR using only CMC slices. Each CMC sliceable spacetime is now represented by a one parameter family of points in N'. If we set $\omega'=\omega|_{N'}$, then our theory of GR is $(N',\omega',0)$. This is a gauge theory. But whereas the gauge orbits of our initial theory, $(N,\omega,0)$, were infinite dimensional, we find that the gauge orbits of $(N',\omega',0)$ are one dimensional. All that remains of the

[129] At one time it was hoped that almost all solutions of the EFE might be CMC sliceable. In recent years it has become clear that this is not true—indeed there is an open set of non-CMC sliceable spacetimes. See Isenberg and Moncrief 1996.

[130] The CMC gauge condition is less attractive in light of the discovery, mentioned in the previous footnote, that this move has nontrivial physical content. Nonetheless, this gauge condition remains an instructive example for my purposes.

original general covariance of GR is the ability to reparameterize the dynamical trajectories which represent the CMC slicings.

(iii) Finally, we gauge fix the time parameterization, by requiring that our dynamical trajectories be parameterized by τ. This is a commitment to use τ as our time. This makes GR into a time dependent Hamiltonian system, $(N',\omega',H(\tau))$.[131] The Hamiltonian $H(\tau)$ is just the volume of the CMC slice of mean curvature τ.[132]

The gauge fixed system $(N',\omega',H(\tau))$ is, of course, equivalent to the original formulation of GR as a gauge theory, in the sense that the dynamical trajectories of both theories correspond to the same set of models of GR (modulo the class of models which are not CMC sliceable). Thus, at the level of formalism, we do not change the physical content of the classical theory by shifting from a traditional gauge theoretic formulation to this CMC gauge fixed formulation, since we do not throw away any models of the theory by doing so.

So long as one is working with GR, one is not likely to make too much of the CMC formulation of GR. Imposing the CMC gauge condition is tantamount to a selection of an absolute time for GR. This seems wildly implausible from the perspective of the classical theory. Thus one is likely to regard the CMC gauge as a useful bit of formalism which holds no interpretative lessons (compare with the discussion of harmonic coordinates in §6.7).

---

[131] See §5.1 of Abraham and Marsden 1978 for a treatment of such systems. As Isham notes, it is quite strange to think of the geometry of a closed universe being governed by a time dependent Hamiltonian: time dependent Hamiltonian systems are usually only employed in modeling physical systems which are driven by some external force (Isham 1991, p. 200). See §5.3.4 of Isham 1991 and §6 of Kuchar 1992 for other reasons to be skeptical about the prospects of a theory of quantum gravity which is based upon the CMC gauge condition.

[132] Notice that H is a nice local observable. Thus, one way to circumvent Torre's theorem is to fix the gauge. Of course, H has a gauge invariant extension to the full phase space of GR which is *not* local. So from one perspective, we have not really gained anything—we have just made the non-local observable "what is the volume of the slice of constant mean curvature τ?" *appear* to be local by restricting attention to CMC slices.

Everything could change, however, when it comes to quantum gravity. The CMC formulation of GR has proved to be extremely useful in investigations of 2+1 quantum gravity (i.e., the theory which results from quantizing the restriction of GR to three spacetime dimensions). This is an interesting testing ground for proposed solutions to the technical and conceptual problems of quantum gravity (see Matschull 1995 for a survey). The EFE in 2+1 dimensions force spacetime to be flat. So 2+1 GR has only a finite number of degrees of freedom. Thus, 2+1 quantum gravity is quantum mechanics rather than quantum field theory. Yet many of the conceptual problems of full quantum gravity must still be faced (although they may appear in a less virulent form). It turns out to be possible to construct perfectly rigorous canonical quantizations of 2+1 GR—although it is far from straightforward to interpret these theories. In an interesting series of papers, Carlip has investigated the canonical quantization of the CMC formulation of 2+1 GR, and compared it to other quantizations (see Carlip 1990, Carlip 1991, Carlip 1994, and Carlip and Nelson 1995). His conclusion is that the quantization of the CMC formulation of 2+1 GR is equivalent to some, but not all, quantizations of 2+1 GR which do not involve gauge fixation (for the point about inequivalence, see especially Carlip 1994). There are several approaches to quantizing 2+1 gravity which lead to equivalent quantum theories, despite their differences from one another.[133] One can view the members of this set as complementing one another—each has its particular technical and conceptual strengths and weaknesses, but taken together they give a complete view of 2+1 quantum gravity (see, e.g., Ashtekar 1994 or Carlip and Nelson 1995). This is often taken to be a very encouraging sign.

One can hope that this situation will carry over to the full 3+1 case. There is, however, good reason to doubt that this will happen. In general, one expects much more diversity from quantizations of infinite dimensional systems than from quantizations of finite dimensional systems. Thus it is quite possible that imposing the CMC gauge on 3+1

[133] Aside from the CMC approach, each of these approaches takes the full phase space of GR as their starting point.

GR will lead to a theory of quantum gravity which cannot be reached by other techniques of quantization.

Thus far, the discussion has been about formalism—the CMC gauge is useful condition to impose on GR because it makes the theory more tractable for the purposes of quantization. But should we take it seriously when it comes time to interpret the theory? That is, should we apply the interpretative techniques of §6.7 and view the CMC gauge as a condition governing which states are physically possible? From the perspective of 2+1 gravity, there is no need to do so—the CMC gauge condition is one technique for quantizing GR among a number of equivalent techniques. It is the only one which involves gauge fixation. So taking the CMC gauge seriously is incompatible with a naive reading of the other approaches to quantization. Thus there is no reason to privilege the CMC gauge approach above these others by assuming that it reflects some deep fact about physical possibility.

The situation may well change, however, when it comes to full 3+1 quantum gravity. If the equivalence between diverse approaches to quantizing gravity does indeed dissolve in the four dimensional case, then it could happen that a quantization based upon the CMC gauge condition leads to empirical predictions which cannot be reproduced by any other approach. If these predictions are borne out, then we will have to regard the CMC gauge as providing a deep insight into the nature of gravity, since it is only by imposing this gauge condition that we can cast GR into a form which leads to an empirically adequate quantum theory.

Let's suppose that this happens. What does it tell us about GR?

The CMC gauge condition amounts to a choice of time for any globally hyperbolic model of GR. The time assigned to any event is just the value of the mean curvature of the CMC slice on which that event lies. Thus there is a preferred slicing of each model into surfaces of simultaneity, and a preferred parameterization of the temporal evolution of the geometry of space. If we follow the approach of §6.7 and view the CMC gauge condition as representing a constraint on physical possibility, then we should adopt a three dimensional interpretation of GR, under which the theory governs the temporal evolution of spatial geometry.

Under this sort of literal interpretation of CMC gauge-fixed GR, spatial geometry evolves in a sort of semi-absolute time. This time is absolute in a couple of ways: it induces a preferred slicing of physical spacetime into surfaces of simultaneity (space); and it has a preferred parameterization. In both respects, it is a departure from our ordinary understanding of time in GR. It is not, however, the absolute time of Newtonian physics. This is because the preferred parameterization is itself a dynamical feature of the theory—time is parameterized by mean curvature, which is governed by dynamics. The rate at which time flows, therefore, depends upon the distribution of matter. Here the contrast between the theories relationalist theories of gravitation, LB2 and BB, may be helpful (see examples 5.6 and 5.8). LB2 may be viewed as arising out of BB via a judicious choice of temporal gauge, analogous to the CMC gauge condition. From the perspective of LB2, time is absolute. Indeed, Lynden-Bell constructs the theory by extending the invariance group of Newtonian physics to include time dependent orientation preserving isometries of Euclidean space without altering the role of absolute time (Lynden-Bell 1995). But from the perspective of BB, there is no preferred parameterization of time. There is, however, a fact of the matter about the rate at which clocks run. But this fact is *dynamical*—"the rate at which a… clock ticks is dynamically determined by the distribution of energy in the entire system" (Smolin 1991, p. 259). When one employs such a clock, one can be deceived into believing that the flow of time is non-dynamical. Just so under a literal interpretation of CMC gauge-fixed GR: time is absolute in that it has a preferred parameterization and induces a preferred slicing of spacetime into space; but differs from the absolute time of Newtonian physics in that it is dynamically determined. Because it falls short of true absoluteness I call it *semi-absolute* time.

So much for the nature of time under a literal interpretation of CMC gauge fixed GR. What about the nature of space? Here we have a choice: three dimensional relationalism or three dimensional Lockean substantivalism. Both strategies support a literal, and hence indeterministic, interpretation of the formulation of GR based upon superspace (see §§7.4 and 7.5). Thus either will support a deterministic interpretation of the CMC version of GR (simply throw away most of the possibilities countenanced under the corresponding interpretations of the full theory).

I am inclined to believe, however, that relationalism has a decisive advantage in this case. This point is made most clearly by returning to the contrast between LB2 and the Newtonian theory of gravitation. It is natural to give LB2 a relationalist reading (see example 6.3). Consider the transition from the configuration space, Q, of Newtonian physics (which is the space of configurations of a set of particles relative to Euclidean space), to the configuration space, Q_0, of LB2 (which is the set of relative configurations—the set of relative distances between the particles). The action of the group of Euclidean isometries, which carries us from Q to Q_0, seems to erase the need for spacetime points—if we are to identify all of the points in Q which represent configurations relative to space which exhibit the same relative configuration, why not just reduce our ontology to particles and their relations and forget about points? It is, of course, *possible* to give a Lockean reading of LB2. That is, one can maintain that space is a real existent whose parts maintain their identity over time, but (somehow) claim that it makes no sense to imagine that *this* group of particles could have been over *there*. This allows one to deny the Leibniz-Clarke counterfactuals, and hence to give a Lockean substantivalist interpretation of LB2 set in T^*Q_0. This position is not logically incoherent. But it is wildly implausible. Indeed, it seems to be exactly the sort of move that substantivalists are always accusing relationalists of attempting: an intellectually dishonest attempt to warp an interpretation to fit a formalism that is manifestly better suited to a rival interpretative strategy. The situation is very similar when we consider the possibility of giving a Lockean substantivalist interpretation of GR set in superspace. Here the action of the group Diff Σ washes out the need for points when we pass from the configuration space Riem Σ to the configuration space Riem Σ/Diff Σ (see Barbour 1994a for this sort of reading).[134]

This is, to some extent, a matter of taste. In any case, it is clear that taking the CMC gauge condition seriously forces an interpretation of GR upon us which is very different from our ordinary one. Thus we have a case where the success a particular solution to the problem of time in quantum gravity would lead us to favor a certain sort of interpretative approach to GR. On the other hand, of course, if there were anyone who was already committed to such a solution, it would lead them to favor quantizing the CMC gauge fixed formulation of GR. Here we have a very clear instance of the interpretative interaction between classical and quantum gravity.

(5) Observables and constraints: Kuchar's maneuver. It is suggested from time to time that the way to introduce time into quantum gravity is to employ observables which do not commute with the constraints (see, e.g., p. 2601 of Unruh and Wald 1989). Kuchar has been the most persistent and explicit champion of this suggestion (see especially pp. 135-42 of Kuchar 1993a; see also pp. 138-40 of Ashtekar and Stachel 1991 and §6 of Kuchar 1992). Kuchar's point of view has recently been taken up by Hajicek (see Hajicek 1996; see also Hajicek and Isham 1996a). To date this is very much a proposal for future research. It is, therefore, quite sketchy in comparison with the approach of the previous example.

Kuchar's proposal for resolving the problem of time comes in two parts. The first involves the sort of program outlined at the end of §4. That is, Kuchar advocates an internal time framework, in which one first identifies a family of functions on the phase space of GR which can serve as time functions. One then uses this structure to rewrite the classical scalar constraint so that it is linear in the momentum conjugate to time. Imposing the quantum version of the scalar constraint is then equivalent to requiring that states satisfy some sort of

[134] Note that the above is not quite a general argument for relationalism—it only works for three dimensional interpretations.

generalization of the Schrödinger equation (see pp. 141-42 of Kuchar 1993a for a sketch of this proposal). This program is extremely difficult to execute—to date very little progress has been made in the case of full quantum gravity (see §5.2 of Isham 1991, §4 of Isham 1993, and §§6-8 of Kuchar 1992 for the problems facing this sort of approach).

The second part of Kuchar's approach to time in quantum gravity requires a revision of the orthodox doctrine which requires that the observables of quantum gravity should commute with the constraints. Part of the rationale here is technical: Kuchar points out that if we work with gauge invariant observables, such as Rovelli's constants of motion, then the expressions for the classical observables in terms of the fundamental variables will typically be very complicated; when we attempt to find operators which correspond to these quantities we will be faced with formidable technical problems, such as factor ordering ambiguity (see Hajicek 1991 and §15 of Kuchar 1992). But there is also an underlying conceptual motivation. Kuchar believes that it is a conceptual error to require that the observables commute with the scalar constraint (see pp. 135-41 of Kuchar 1993a; see Belot and Earman 1996b for critical discussion). His rationale is that the scalar constraint is responsible for time evolution, so that to require that the observables commute with the constraint is to opt for a conceptual scheme in which physics is timeless. From this perspective, the problem of time in quantum gravity is only to be expected—it was built in to the framework from the beginning by our decision to treat the vector and the scalar constraints on the same footing.

Kuchar himself is a non-Lockean four dimensional substantivalist about GR.[135] I believe that this interpretative commitment plays a crucial role in justifying his position concerning the observables of quantum gravity. I am going to present a justification along these lines—I will attempt to explain why it is reasonable to hold that the observables of quantum gravity do not commute with all of the constraints if one is a substantivalist about GR. I christen this *Kuchar's maneuver*. I do, in fact, believe that this maneuver is Kuchar's.

[135] This claim is based upon private communication as well as his published work.

But I will not insist on this point here, since its justification would be quite messy.[136] It is enough, for my purposes, that the maneuver exists, whether or not it is actually due to Kuchar.

Suppose that you are committed to the sort of predictable, non-Lockean, four dimensional substantivalism which was sketched in §7.5. That is, you hold: (i) that the possible worlds described by vacuum GR consist of physical spacetime points which stand in geometric relations to one another; (ii) that because these physical spacetime points are existents, it is possible that the geometry of the world could have been just as it is, but with the geometric properties shared out differently among the points; (iii) such a world would be empirically indistinguishable from the actual world. From (ii), we see that you hold that the beables of GR do not commute with either the vector or the scalar constraint. From (iii), however, it follows that you believe that the observables commute with both constraints.

Does this commit you to a particular position on the question of whether or not the observables of quantum gravity should commute with the quantum constraints? I claim that it does not. In order to approach this question, we need to address an issue which has been present but dormant since the §2: do the observables of a quantum theory correspond to the beables or the observables of the classical theory upon which it is based?

It might be thought that the issue would be settled by one's attitude towards operationalism. Isn't the decision to employ quantum observables which correspond to classical observables rather than beables a decision to ignore ontological distinctions which are not accessible to empirical determination? And won't our attitude towards this decision be determined by our attitude towards operationalism or positivism in general? This line of thought is not quite on the mark. Notice that the question at hand is the nature of the

[136] To mention just one problem: at times Kuchar maintains that the classical observables should commute with all of the constraints (p. 139 of Ashtekar and Stachel 1991); elsewhere, he insists that the observables of GR should commute with the vector constraint alone (see pp. 135-41 of Kuchar 1993a). I believe that, despite appearances, Kuchar is not inconsistent. But it requires quite a bit of work to rescue him, since one must read him as using the word 'observable' in a number of distinct senses.

quantum observables—we have already settled on a set of quantum states.[137] That is, we have already decided upon a set of possible ways for the world to be, and we are now trying to decide upon a model for the way in which these ontological possibilities can be made manifest. And here we may, in good conscience, restrict our attention to physical quantities whose classical counterparts are observables rather than beables if we so choose. This is especially true in light of the fact that our conception of classical observables was by no means operationalistic—observable meant "observable in principle" rather than "observable in practice."

In the case at hand, there is actually a consideration of physics which mitigates in favor of requiring the observables of the quantum theory to correspond to classical observables rather than to beables. For if we have a quantum observable which fails to commute with the quantum scalar constraint , then we have $\psi - \psi \neq 0$ for $\psi \in V_{phys}$. But the quantum constraint annihilates physical states, so $\psi = 0$. Thus we find that $\psi \neq 0$. This means that $\psi \notin V_{phys}$. This result may be problematic when it comes time to formulate a measurement theory, since, naively, it says that a measurement knocks the state out of the space of physical states.[138]

This is not by any means a decisive consideration. Quantum gravity, at least in its role as a theory of the entire universe, will require an innovative measurement theory since it will be necessary to include the observer in the model (see Smolin 1991 for discussion). It seems highly unlikely that such a measurement theory will include a collapse postulate. So

---

[137] It may seem suspicious to claim that the choice of a vector space of states is independent of the choice of a set of operators on that space. In ordinary quantum mechanics, after all, we select the space of states by looking for a space which carries a representation of the algebra of observables. Recall, however, that under the schema of §2, the space of states is selected because it carries a representation of the algebra of fundamental variables. These will not, however, necessarily be observables (quantities like the value of the metric tensor at a point are not gauge invariant). Thus the question of the nature of the observables is distinct from the question of the size of the space of states.

[138] See p. 2601 of Unruh and Wald 1989 for this sort of objection to schemes which permit quantum observables which fail to commute with the constraints. Kuchar is, of course, well aware of this problem. See p. 232 of Kuchar 1992 and p. 142 of Kuchar 1993a.

it would seem that we need not worry too much about the fact that quantum observables which fail to commute with the constraints expel states from the space of physical states.

I contend that neither philosophy nor physics singles out a determinate answer to the question of whether the observables of quantum gravity correspond to the observables or the beables of GR. Up until now, this question has been neither obvious nor important to physicists working on quantizing classical theories. One has hunches and plausibility arguments, but no decisive considerations. Since neither theory nor practice singles out an answer, one is free to legislate for oneself.

In the case at hand, it means that substantivalists are free to stipulate whether or not the quantum observables commute with the constraints. Kuchar's maneuver requires that the quantum observables commute with the vector but not with the scalar constraint. The justification is as follows. According to the sort of substantivalism which Kuchar endorses, the observables commute with both vector and the scalar constraints, while the beables commute with neither. There is, nonetheless, an important difference between the vector and the scalar constraint. Consider three points, x, x', and x", in the phase space of GR. Suppose that x' is related to x by a gauge transformation generated by the vector constraint, while x" is related to x by a gauge transformation generated by the scalar constraint. Kuchar holds that these points represent distinct physical possibilities—different ways that the points of some slice of physical spacetime could be assigned geometrical properties. But the possibilities represented by x and x' are qualitatively identical—they represent the same spatial geometry, but distribute the underlying geometrical properties over the points of space in different ways. x and x", on the other hand, represent distinct spatial geometries. Thus the vector constraint generates gauge transformations between qualitatively identical situations while the scalar constraint generates gauge transformations between qualitatively distinct physical situations. Under Kuchar's maneuver, when it comes time to construct quantum observables, we only care about the second kind of difference. That is, our quantum theory should not give us observational resources which derive from a distinction between qualitatively identical classical states.

Kuchar's maneuver is very subtle. But it is worth exploring since it opens up an avenue of research on quantum gravity which (Kuchar hopes) will lead to a theory which is considerably easier to make sense of than a theory which results from the standard approach. And here we see a commitment to a certain interpretative doctrine about GR making a substantive difference to how one approaches quantum gravity. The doctrine that the quantum observables should fail to commute with (any of) the constraints finds its easiest justification in an appeal to an interpretation of GR under which the beables do not commute with the constraints, combined with a decision that the quantum observables should correspond to classical beables. Indeed, I cannot not think of any other justification for this sort of approach to quantum gravity. Thus we have an example of an approach to quantum gravity which is readily available only to those who deny that the beables of the classical theory commute with the constraints. In fact, of course, the only such classical interpretations are: non-Lockean four dimensional substantivalism; and various three dimensional interpretations. Since the latter have very few adherents, it is safe to say that we have discovered a connection between non-Lockean substantivalism and a certain approach to quantum gravity. The former does not make the latter mandatory—one can always opt for quantum observables which correspond to classical observables, and hence commute with all of the constraints. But substantivalism is, modulo three dimensional interpretations, a necessary condition for loyalty to the sort of approach to quantum gravity that Kuchar advocates. Substantivalists will have good reason to celebrate, should this approach produce a viable theory of quantum gravity—since their interpretative position will have been vindicated by fundamental physics.

This completes the argument which has extended over the entire second Part of this dissertation. In Chapter 6, I argued that gauge theories admit multiple interpretations. In particular, they admit literal, indeterministic interpretations under which the beables do not commute with the constraints, and gauge invariant, deterministic, interpretations under which the beables do commute with the constraints. In Chapter 7, I argued that attending to the fact that GR is a gauge theory should make us worry about what sort of interpretation we give it—since it will be indeterministic under literal interpretations and deterministic under gauge invariant interceptions. I classified the interpretative options for GR in terms of the commutation with the vector and scalar constraint of their respective beables. Finally, in the present Chapter, I have shown that there is an interpretative interplay between classical and quantum gravity. Attempts to solve the problem of time in quantum gravity are often associated with one or another interpretative option for GR. As a consequence, the attractiveness of interpretative strategies for dealing with the spacetime of GR hangs in the balance. If a given strategy for dealing with the problem of time appears to be the only viable approach, then we will favor any interpretations of GR which are associated with it—even in this means believing in a preferred notion of simultaneity, or a variety of substantivalism under which GR indeterministic. Conversely, if we are antecedently committed to a given interpretation of GR, this will affect our attitude towards various proposed solutions to the problem of time.

Furthermore, notice the structure of the present example: non-Lockean substantivalism suggests an approach to quantum gravity which is incompatible with relationalism. This is driven by a difference of opinion between relationalism and non-Lockean substantivalism regarding the question of whether or not the beables of GR commute with the constraints. Notice that Lockean substantivalism, whose beables and observables have exactly the same commutation properties as those of relationalism, is just as incompatible with Kuchar's proposal as relationalism is. I believe that this is not an isolated example—I expect that where relationalism and substantivalism are associated with distinct approaches to quantum gravity, this will be a consequence of a disagreement concerning the commutation properties of the classical beables. If this is right, then Lockean substantivalists will always side with relationalists when it comes to disagreements in quantum gravity which are driven by interpretative differences concerning GR.

For this reason, I regard Lockean substantivalism as an uninteresting option. It requires some work to establish that it is possible for a substantivalist to deny the Leibniz-Clarke counterfactuals. And if one succeeds in doing so, one’s reward is to occupy a conceptual space already occupied by relationalism. This is to turn one’s back on the hope, offered by quantum gravity, that the substantival-relational debate can be resuscitated by being brought back into contact with physics.

AFTERWORD

Before bringing this dissertation to a close, I would like to take the time to recapitulate my central argument, and to offer some comments about the directions in which this work points.

I began in Chapter 1 by attempting to situate the substantival-relational debate. I noted that the way in which physicists view the debate diverges wildly from the way in which philosophers view it. Many philosophers believe that the debate, once vital and closely intertwined with physics, has in recent decades grown moribund and lost all relevance to physics. At the same time, there is a wide consensus among philosophers that the substantival-relational debate isn't really much of a debate anymore—the combination of relativistic physics and the new philosophy of space and time of the 1970's having led to a decisive victory for substantivalism. On the other hand, physicists are more evenly divided into substantivalists and relationalists. Furthermore, there is a vocal faction of physicists working on quantum gravity who claim that there is an intimate connection between interpretative issues in classical and quantum gravity. In particular, they claim that some of their disagreements with one another concerning the best way to proceed in the face of the formidable technical and conceptual obstacles which stand in the way of their research program can be traced to differences of opinion concerning the ontological status of the spacetime points of GR. I set myself the project of analyzing this interpretative interrelationship between classical and quantum gravity.

My first task was to talk my way out of the received philosophical view that relationalism is not a viable strategy for interpreting GR. I began this task by arguing that the connection which many philosophers presume to exists between substantivalism and scientific realism is in fact nonexistent. Many discussions of the substantival-relational debate presuppose that scientific realism entails substantivalism, while anti-realism of an empiricist variety entails relationalism. I argued that this is not the case—the two issues are logically independent. This effectively deactivates a concern which has motivated many philosophers to reject relationalism. I went on to redescribe the substantival-relational debate as being concerned with interpreting GR, conceived of as one false theory among many. This description derives from a view of contemporary theoretical physics as consisting of a web of interrelated theories, each of which is known to be false, and yet to have a domain of empirical adequacy. I contend that one of the central tasks of philosophy of physics is to aid in the project of articulating the content of physical theory. This often means attempting to understand what a given physical theory, taken in isolation, is telling us about the world. For philosophers of physics, this usually involves interpreting the theory—i.e., saying what way(s) the world would have to be in order for the theory to be true. I contend that the substantival-relational debate should be conceived of as being part of this interpretative enterprise. As such, it is concerned with articulating the content of science, and hence is prior to, and independent of, the realism-anti-realism debate (which is concerned with the question of what epistemological attitude towards the content of science is warranted).

I also argued that it is an important part of the project of articulating the content of physical science to look at theories in relation to one another, as well as in isolation. This is, I believe, a task which is often overlooked by philosophers of physics, except in very special cases such as the measurement problem in quantum mechanics and the reducibility of thermodynamics to statistical mechanics. There is, however, much interesting physics and philosophy of physics to be found in the links between theories. I conceive of my project in this dissertation as providing a detailed example of the way in which consideration of the technical and conceptual links between two theories can set up a correspondence between the interpretative problems of the theories.

This was all by way of setting up the discussion of the body of the dissertation. In Part I, I criticized the received view of the substantival-relational debate. I began by considering substantivalism and relationalism about the space of Newtonian physics. Substantivalism was construed as the doctrine that space has parts which endure through time and maintain their identities, thus allowing a definition of absolute motion in terms of the relations between bodies and the parts of space. Relationalism was construed as the doctrine that space, although existent, should not be thought of as being comprised of genidentical parts—rather it is the framework of possible spatial relations between bodies. I found that neither of these doctrines was entirely attractive as a strategy for interpreting Newtonian physics. But I argued that the objections against relationalism about space upon which the received view is founded are either chimerical (depending upon non-existent necessary connections between substantivalism and realism on the one hand, or between relationalism and empiricism on the other) or very weak (depending upon a conservative assessment of the ideology available to relationalists).

Both interpretative approaches are strengthened by a transition to four dimensions. Here substantivalists hold that the parts of spacetime are ephemeral, but nonetheless genuine, existents; and that the spatiotemporal relations between material events should be thought of as deriving from the spatiotemporal relations between the spacetime points occupied by events. Spacetime relationalists, whose position is often overlooked in philosophical discussion of the substantival-relational debate, view spacetime as being the geometry of possible spatiotemporal relations between material events; spacetime points play no role in their ontology. I argued that spacetime relationalism is at least as promising as spacetime substantivalism as a strategy for interpreting Newtonian and relativistic physics.

Throughout this discussion, the Leibniz-Clarke counterfactuals served as a criterion for separating relationalists and substantivalists. Substantivalists, but not relationalists, believe that the parts of space(time) are individuals in their own right, and are thus bearers of properties. Thus one expects that substantivalists, but not relationalists, will affirm the existence of distinct but qualitatively identical worlds, which differ only as to which point of space(time) plays which role (of being occupied by *this* bit of matter, of having *this* geometrical property). I acknowledged that this rule of thumb is not universally applicable—it is possible to follow Locke in proclaiming oneself a substantivalist while denying the truth of the Leibniz-Clarke counterfactuals. But this Lockean position can be secured only at a price—one must either work very hard to show that the falsehood/senselessness of the Leibniz-Clarke counterfactuals follows naturally from a general account of modality and from the structure of our modal intuitions; or stipulate that spacetime points are special kinds of objects whose modal semantics is different from that of ordinary existents. I am inclined to believe that either price is more than substantivalism is worth.

Thus I began the second Part with the conviction that there are a number of viable interpretations of GR, and with the presumption that the Leibniz-Clarke counterfactuals provide a shibboleth for distinguishing substantivalists from relationalists. In the first Part, I had searched for *physical* reasons for preferring one or another of these interpretative strategies. I did not find them. Following DiSalle—"most of us do not have accepted metaphysical principles which we take more seriously than physics and from which we can hope to derive knowledge of space and time" (DiSalle 1995, p. 285)—my attitude was that any reason of physics which emerged from my discussion of the relation between classical and quantum gravity in Part II would be sufficient to trump the purely metaphysical reasons which emerged from the discussion of Part I.

The discussion of Part II centered around the question of the relevance to interpretative questions of the general covariance of GR. In the philosophical literature, this question received its first clear formulation in the hole argument of Earman and Norton 1987. I brought this question into sharper focus by setting the general covariance of GR in the more general context of gauge freedom. I spent the first two Chapters of Part II in reviewing the formalism of gauge theories, and the interpretative problems which arise out of this formalism. In particular, I argued that indeterminism is a threat whenever we are dealing with a gauge theory, since such theories are *always* indeterministic under their most obvious interpretation—i.e., under any literal interpretation according to which there is a bijection between the mathematical structure of the points of phase space and the physical ontology of the world which the theory represents.

In Chapter 7, I went on to consider the standard Hamiltonian formulation of GR as a gauge theory. I showed that the constraints which generate gauge transformations are useful in classifying the interpretative possibilities for GR. The main conclusion here was that, among four dimensional interpretations, non-Lockean substantivalism is the only indeterministic interpretation of GR (this is the hole argument). I do not, however, regard this as being a decisive defeat for this brand of substantivalism. The discussion of Chapter 8 shows how stances on interpretative questions concerning GR can impinge upon the ongoing project of constructing a theory of quantum gravity. There is a sort of interpretative interplay between classical and quantum gravity—adopting a position on the interpretation of either theory will affect one's attitudes towards strategies for interpreting the other. In particular, the same feature which makes non-Lockean substantivalism indeterministic (i.e., the fact that its beables fail to commute with the constraints), may also make it desirable, since it opens up an option for conceiving of the role of time in quantum gravity which is unavailable under other interpretations of the classical theory.

Thus we see that the hole argument, far from being an idle fancy of philosophers who have lost touch with physics, is in fact part of a knot of conceptual problems at the center of research on quantum gravity. It is in this sort of interpretative interaction between classical and quantum gravity that I see the promise of an era of renewed relevance and vitality for the substantival-relational debate.

This completes my summary of the central argument of the dissertation. In conclusion, I would like to offer a few comments about topics for further research which are suggested by this dissertation.

Most narrowly, the projects initiated in Chapters 6, 7, and 8 are far from completion. Philosophers of physics have paid too little attention to gauge theories and their interpretation. The fact that there is so little discussion of the interpretative problems of classical electromagnetism is the most obvious symptom of this shortcoming. Here we need a more fine-grained, general, account of the interpretative options available for gauge theories, as well as detailed work on specific theories. GR is, perhaps, the most important special case. And here there is a great deal of work to be done. Some of this work involves an exploration of the interpretative interaction between classical and quantum gravity. Examples 8.4 and 8.5 are, I am confident, the tip of the iceberg. I expect that many—if not most—of the proposed solutions to the problem of time are related to some sort of interpretative move concerning the nature of the spacetime of GR.

More broadly, as I have commented a number of times, I believe that philosophy of physics suffers from a dearth of studies of intertheoretic relations and their significance (Batterman 1995 and Rohrlich 1996 are important recent exceptions).[139] Philosophers of physics tend to be philosophers of a single theory, X, and to face the interpretative problems of X by invoking the fiction that X is true. This often leads to a fruitful and interesting discussion of the interpretative problems of X, which tells us quite a bit about what the world would be like were X true. But, of course, as a matter of fact X is never true—no theory is both fully relativistic and fully quantum. This has an important consequence which is often overlooked. There is another way of gleaning information from X about what the world is like, beyond pretending that X is true and that the world is as it would have to be for this to be the case. One can also look at the way in which X is related to some other theory Y which describes phenomena of our world which are related to those described by X, and by looking at how the descriptions of X and Y mesh, one can discover quite a bit about the way the world is. This is the sort of thing that we saw in Chapter 8—not only did we see an interpretative interplay between classical and quantum gravity, but also examples

[139] As mentioned above, there are exceptions to this trend in special cases, such as the measurement problem in quantum mechanics or the problem of the direction of time in statistical mechanics, where the outstanding interpretative problem of a theory is really a problem about the relation between theories. I maintain that such cases are indeed the exception rather than the rule.

taken from ordinary quantum mechanics and quantum field theory where the ambiguity inherent in quantization sets up an interpretative interaction between two theories.[140] In terms of my metaphor of our web of physical theories: if one looks only at the nodes of the web, one gains only partial information; in order to grasp the full content of the fact that this web is useful for describing our world, one must also look at the strands which join these vertices. If one wants to survey the shape of an object caught in a net, it is useful to note how the threads bulge, as well as where they meet. This strikes me as being highly suggestive—I suspect that the sort of program which was executed in this dissertation could be executed for any number of pairs of theories which stand in some sort of limiting relation to one another. By considering theories in relation to one another rather than in isolation, philosophers of physics can expect to achieve a richer appreciation of the content of our physical theories.

Finally, and most broadly, I believe that the approach developed in this dissertation has implications for general philosophy of science. The central antitheses of recent philosophy of science are realism-anti-realism and rationality-relativism. In discussions of both issues, one finds the same sort of concentration on physical theories taken in isolation from one another as one does in philosophy of physics, with similarly deleterious consequences.

The realism-anti-realism issue concerns the question of what epistemic attitude is warranted towards the entities and laws mentioned in our best theories. Here it is always supposed that there *is* a single best theory describing the phenomenon of a given domain, and that to take the entities of physics seriously, or with a grain of salt, is to believe, or not,

---

[140] Notice that the two theories that stand in the relation of mutual interpretative dependence induced by quantization may, but need not be, the classical theory at hand and its quantization. In examples 8.2, 8.4, and 8.5 we see an interaction between a theory and its quantization. But in example 8.3, we see that certain approaches to managing the open-endedness of quantization can set up an interpretative interrelation between two theories which stand in some relation other than classical-quantum—such as general relativistic-special relativistic or thermodynamical-mechanical. Establishing a link between two theories, as is done when one quantizes, may sometimes require adjustments in the fabric of the web of theories in places far removed from the link under construction.

in the entities postulated by this unique best theory. One considers theories in relation to one another only when one asks a question such as: When we look at a historical sequence of physical theories governing a single domain—such as Newtonian gravity→GR→quantum gravity—do we find that later theories are closer to the truth in some sense, and that later theories refer to the same entities as earlier theories do?[141]

I object to this approach on the grounds that it ignores most of the richness of scientific practice. It is fine to honor quantum gravity, or whatever, with the title "our best physical theory." But it is a mistake to go on to disregard the important fact that there are other theories of gravitation which are fruitfully employed by physicists. Surely the fact that these other theories are part of a network of theories which we use to model the world contains some information over and above that contained in the fact that our best theory is empirically adequate in its domain. And surely this sort of surplus information should play a role in any philosophical account of metaphysical and epistemological ramifications of the teachings of science.[142]

Realists ask whether we should believe in the entities postulated by our best physical theories. Should we, for instance, believe in spacetime? Well, our best theory of gravitation, GR, talks about spacetime. This is supposed to be a reason to believe in spacetime (although, as we have seen, the question of the existence of the parts of spacetime is an open one, even for realists). Some of our other physical theories, however, suggest that the spacetime continuum may not exist (quantum gravity, high energy physics). Similar points can be made about the structure of spacetime—our best theory of gravitation requires a curved Lorentzian spacetime; our best theories of molecules require a Newtonian spacetime; our best theories of continua can be set in something like Leibnizean spacetime (Wilson 1993). Realists ask physicists to tell them what sort of ontology to believe in. The answer

[141] See, e.g., Laudan 1984 and Leplin 1984.

[142] I will not accept the objection that there is no information contained in GR which is not already contained in quantum gravity, since the former arises as the classical limit of the latter. Attention to the details of limiting relations between theories reveals that, in addition to being complex to a degree that is in general beyond our ability to analyze, they also need not preserve information. See Berry 1994.

that they get will depend on which physicist they are talking to. What then, should realists do? I suggest that they should abandon their favorite questions—Do electrons exist? Does the ether?—which presuppose that theoretical physics is monolithic, and attempt to articulate an epistemological doctrine appropriate for the actual complex structure of physics.[143]

Discussions of the historical relativity of scientific knowledge also tend to proceed from the assumption that a the scientific community has one theory of a given type on its books at any given time—Newtonian gravity was replaced by GR, which will in turn be replaced by quantum gravity. According to Kuhn 1970, communication is impossible between adherents of distinct theories in this sequence—if I work within the Newtonian paradigm, then I will be literally unable to understand the what Einstein means by 'mass.' Relativism of one sort or another is supposed to follow easily from this sort of incommensurability thesis. More recent treatments of incommensurability tend to be both more sophisticated and less directly tied to relativism. Typically, they allow for various sorts of partial communication between proponents of distinct theories, via partial translation, bilingualism, or pidgins (see, e.g., Biagioli 1990, Buchwald 1992, Galison 1995, and Kuhn 1991). But the fact remains that the transition from, e.g., Newtonian gravity to GR is supposed to create a linguistic gulf which it is, if not impossible, then quite difficult to bridge.

I suggest that this view is only plausible so long as one has in mind the following picture: the history of science is the history of a community of inquirers who pursue a given theory at one time, only to replace it by another at a later time. On this view, it is perfectly

[143] Here is Cassirer, on the subject of making sense of the question of the reality of atoms in light of the historical succession of atomic theories: "If, in this progression, we can move in a single direction and be sure of it, we have in this direction and in this alone the criterion of truth. Thus we need not find an answer to the question of what the atom actually and ultimately is, as long as we are sure that the various forms of determination given us by progressing experimental and theoretical investigation can dovetail and mutually supplement one another. As long as this empirical interconnection is safeguarded, and as long as we are convinced that the progression of atomic models is guided not by mood or caprice but by a definite rule, the atom has all the reality of which it is capable"; Cassirer 1956, p. 150. I would hope that something similar can be said of the network of physical theories which are accepted at any given time. See also the discussion of 'harmonious integration' in Hacking 1996.

sensible to inquire after the conditions under which communication is possible between partisans of successive theories. But this picture of science is untenable, at least as applied to theoretical physics. The history of theoretical physics is not just the history of the series of 'best' theories. Contemporary theoretical physics, as a network of theories, includes classical mechanics and non-relativistic quantum mechanics, both of which have been superseded by more fundamental theories. Furthermore, most physicists are conversant with a number of theories which describe a given domain of phenomena, and can move back and forth between them, apparently without effort. In Wald 1994, for example, one finds considerations and techniques drawn from classical mechanics, quantum mechanics, quantum field theory, GR, quantum gravity, and thermodynamics as the need arises. These are blended smoothly together to support the central theses of the book. Wald shows none of the historians' discomfort in moving between conceptual schemes. Any account of science which purports to engage with the details of practice must come to terms with the fact that contemporary theoretical physics is comprised not of a handful of current 'best' theories, but of a network of coexisting theories, old and new.

APPENDIX

APPENDIX

Modal Justifications of Lockean Substantivalism

Lockean substantivalists deny the Leibniz-Clarke counterfactuals. In the case of GR, this means that they deny that there could exist two qualitatively identical but distinct possible worlds, sharing the same spatiotemporal geometry and set of spacetime points, but differing as to how the geometrical properties are shared out over the points of physical spacetime.

In the body of the dissertation, I argued that Lockean substantivalism is an unattractive option—it occupies a conceptual space already occupied by traditional relationalism. This means that in denying the Leibniz-Clarke counterfactuals, Lockean substantivalists are foregoing the opportunity to advocate a version of substantivalism which offers the hope of differing from relationalism in some way which is relevant to physics (see §§7.6 and 8.5). This is to consign the substantival-relational debate to a thoroughly avoidable status of irrelevance.

Nonetheless, I concede that Lockean substantivalism is a genuine interpretative possibility (see §3.6). At the very least, one can simply stipulate that the peculiar nature of spacetime points does not support the Leibniz-Clarke counterfactuals. To many, this will seem to be about as virtuous as theft. In this Appendix, I will undertake to show that theft is the only available option—those who have attempted to justify Lockean substantivalism via honest toil have failed. In the part A, I will argue that the essentialist approach to justifying Lockean substantivalism fails because one cannot establish that the metric properties of spacetime points are essential properties. In the part B, I will argue that the counterpart theoretic approach founders because it does not guarantee that the Leibniz-Clarke counterfactuals are indeed false.

A

The essentialist approach is supposed to block the conclusion of the hole argument: it is supposed to show that one can be a respectable substantivalist and still maintain that GR is deterministic (see Maudlin 1989 and Maudlin 1990; see also Bartels 1994).[144] The key claim is that the geometric properties of the spacetime points of GR are essential properties.[145] Thus it makes no sense to consider a world which has the same spacetime points as this one, but in which the points have different properties than the do in our world. It follows immediately that the Leibniz-Clarke counterfactuals fail, and that Lockean substantivalism is a legitimate strategy for interpreting GR.

There are a number of well known objections to this maneuver (see §5 of Butterfield 1989b, §9.13 of Earman 1989, and pp. 62-63 of Norton 1989). The most important runs as follows. One of the most striking features of GR is that it portrays the geometry of spacetime as being dynamical—as depending upon the distribution of matter and energy. But this is just to say that if I had been heavier, the curvature of spacetime hereabouts would have been greater. Which is to say that *this* point could have had geometrical properties other than those which it in fact has. Therefore essentialism is untenable, since it holds that such straightforward assertions are "self-contradictory" (Earman 1989, p. 201).

There is a standard essentialist rebuttal to this objection (see p. 90 of Maudlin 1989, pp. 550-51 of Maudlin 1990, or p. 26 of Bartels 1994). It runs as follows: it is true that we admit the possibility that Belot could have been heavier so that spacetime would have been more curved; but this is not to admit the possibility that *this* spacetime point could have had

[144] I set aside consideration of the non-relativistic case, where I believe that this essentialist maneuver is effective.

[145] For our purposes, it is sufficient to take the geometric properties of points to be exhausted by distances between pairs of points—other geometric properties, such as curvature, can be logically constructed out of these.

geometric properties other than the ones which it actually has; rather, it is to admit the existence of a possible world distinct from our own, whose spacetime is made up of points which are distinct from the spacetime points of the actual world, but similar enough to count as counterparts of actual spacetime points; the counterfactual is true not because *this* spacetime point could have had geometric properties other than those it has, but because there could have been a very similar spacetime point with slightly different geometric properties. That is: essentialists help themselves to a counterpart-theoretic construal of modality whenever they have to handle awkward counterfactuals which appear to imply that geometric properties are not essential.

Most commentators are not impressed with this stratagem for rescuing essentialism. As Brighouse 1994 notes, essentialism and counterpart theory are strange bedfellows—how can X be a counterpart of Y if Y does not have all of the essential properties of X? I believe that this is an effective criticism of the essentialist ploy. But I also believe that to leave it at that is to go much too easily on the essentialists. As it stands the dialectic is this: (i) the essentialists stake out their ground; (ii) it is claimed that their position is untenable, since the dynamic nature of the spatiotemporal geometry of GR is incompatible with the claim that the geometric properties of spacetime points are essential; (iii) essentialists show how they can finesse this objection; (iv) it is pointed out that the essentialist reply is unattractive. I contend, however, that the strength of the original objection (ii) has been underestimated by its proponents. As an objection against the coherence of essentialism, it can indeed be sidestepped by making the moves suggested by Maudlin and Bartels. But the objection can also be seen as an objection against the plausibility of the claim that geometric properties are essential. And in this form, I claim, the objection is conclusive. For, consider how the essentialist thesis could be justified. What Maudlin and Bartels offer are arguments in favor of the claim that some properties of spacetime points must be essential—otherwise we would be committed to the (supposedly absurd) existence of bare particulars. Even if we accept this sort of argument, we are still a long way from being able to show that a given property of spacetime points is essential. How can we establish which properties are essential? Here, I suggest, we have nothing to go on except intuitions ("If we can't imagine a possible world in which Nixon doesn't have a certain property, then it's a necessary condition of someone being Nixon;" Kripke 1980, p. 46). Thus the fact that everyone agrees that if I had been heavier, spacetime would have been more curved counts as strong evidence that geometric properties are *not* essential properties of spacetime points. The essentialist rebuttal is no help here. It does no good to show that our modal intuitions can be overridden if we accept a hybrid account of modality, unless we have some independent reason for identifying the geometric properties of points as being essential. But we could never have such a reason, since our modal intuitions are our *only* guides in constructing a list of essential properties.[146]

B

1. Introduction.

Recall from §3.6 that Butterfield and Brighouse propose to use Lewis's counterpart theory to justify Lockean substantivalism (see Butterfield 1987, Butterfield 1989a, Butterfield 1989b, Brighouse 1994, and Brighouse 1995). In Lewis's scheme, no entity exists in more than one possible world (see Lewis 1973 and Lewis 1986). So there can not be a possible world which has the same spacetime points and geometry as the actual world but in which individual spacetime points have geometric properties different from those that they have in the actual world. This is a trivial result however, and cannot justify Lockean substantivalism. According to counterpart theory, I do not exist in any other possible world.

[146] It might be suggested that the hole argument itself provides reason to overrule our modal intuitions. As should be clear by now, I think it is considerably more interesting to stand by our original intuitions, and to let substantivalism stand or fall on the question of whether or not non-Lockean substantivalism leads to a viable theory of quantum gravity. In any case, if one is going to adopt this line, one may as well save a lot of trouble by following Maidens in simply stipulating that the Leibniz-Clarke counterfactuals are false, rather than taking a lengthy detour through philosophical accounts of modality.

But it is still true that I might have had other properties than those I do have, and this is made true by the fact that there is someone in another possible world who is similar enough to me to count as my counterpart, who has those other properties in his world. Similarly, although there is no other world with the same spacetime points as the actual world, there might be another world with spacetime points which are similar to those of our world, but which are such that *this* point and its counterpart have distinct geometric properties. If this obtains, then the counterpart theoretic version of Lockean substantivalism fails, since the geometry of the world could have been just as it is, while the spacetime points could have had distinct properties.

How do we decide if the counterpart-theoretic Leibniz-Clarke counterfactuals are true? We have to find two possible worlds W and W' with existent spacetime points and identical geometries, and ask whether there is any spacetime point of W whose counterpart in W' does not share all of its geometric properties. But this will *always* be the case. It is qualitative similarity, and qualitative similarity alone, which determines whether two entities are counterparts of one another. But the degree of similarity which is required varies from context to context. In some contexts, any two things of the same natural kind will be counterparts of one another (see pp. 231-32 of Lewis 1986; more on this point below). So there is always some context in which the Leibniz-Clarke counterfactuals will be true—since there are some contexts in which any two spacetime points are counterparts of one another. Thus Butterfield and Brighouse cannot hope to establish a fully Lockean substantivalism according to which the Leibniz-Clarke counterfactuals are false. The best that they can hope for is to show that the Leibniz-Clarke counterfactuals are false in any conversational context which is sufficiently strict to be of relevance to physics. Thus they hope to show that the hole argument fails because discussion of determinism necessarily involves a standard of qualitative similarity which is sufficiently strict to make the Leibniz-Clarke counterfactuals come out false. To this end, they must show that in such a context, two points are counterparts iff they are qualitatively identical.

Brighouse's strategy is to argue that a theory can never be ruled to indeterministic on the basis of the existence of two possible worlds W and W' which are qualitatively identical (see §2.4 of Brighouse 1995). The core of her argument is the claim that if we have two qualitatively identical but distinct worlds, two worlds which differ *haecceitistically*, then this is a difference which can be of no interest to physics: "the haecceitistic features of the qualitatively isomorphic histories simply don't feature into the way the physical states of those worlds evolve. Haecceitistic differences between objects make no difference to the way objects behave" (Brighouse 1995, p. 26). Physical theory describes the evolution of qualitative priorities, so there is no physically relevant difference between the worlds—omniscient scientists living in these two worlds would develop identical accounts of the laws governing their worlds, and of the initial conditions of their respective worlds. Which is just to say that there is no scientifically interesting difference between the worlds. Since determinism is the paradigm of a scientifically interesting concept, it follows that the existence of two such worlds cannot be sufficient to make a theory indeterministic. This is a consistent and prima facie attractive response to the hole argument. But I claim that it founders once the interpretative interaction between classical and quantum gravity is taken into account. In essence, Brighouse wants to identify determinism with what I have called predictability (see §6.5). To the extent that our disagreement is over the meaning of the word 'deterministic,' it is, of course, uninteresting. But something more is at stake here. Brighouse believes, where I do not, that it is observables rather than beables which matter to physics. She holds that a difference between two sates which are ontologically distinct (different beables) but indistinguishable even in principle (same observables) can be of no interest to physics. I believe that example 8.5 shows that Brighouse's position is unattractive—quantization can render unobservable beables very interesting indeed.

Butterfield, on the other hand, attempts to apply Lewis's account of determinism to the special case of the hole argument, and to show that the Leibniz-Clarke counterfactuals come out false under the appropriate construal. The remainder of this Appendix is devoted to showing that this approach, too, is untenable. My argument involves a detailed consideration of Lewis's proposal. In the remainder of this section, I attempt to motivate Lewis's approach. Then, in §2, I discuss some points of methodology; in §3 I present Lewis's definition. In §4 I will present three counterexamples. The first shows that Lewis's definition is too lenient; the second and the third that the obvious modifications of it are likewise unacceptable. I conclude that Lewis's account does not contain the resources required to block the hole argument.

The idea behind Lewis's approach is roughly as follows. Our analysandum is be 'deterministic world.' Our intuitive idea is that the possible world W is deterministic iff there is only one physically possible future for any initial segment S of W. A very rough first try at translating this into the language of counterpart theory yields

Definition 0: a world W is deterministic if, whenever S is an initial segment of W and W' is a world which is physically possible with respect to W with an initial segment S' which is a counterpart of S, then W' is a counterpart of W.[147]

This isn't a very good explication of our commonsense notion of determinism. One of the virtues of Lewis's theory is that it makes the truth values of counterfactual statements depend radically on their context: in different contexts, different counterpart relations are appropriate for evaluating a given counterfactual statement. But our judgments about whether or not a theory or a world are deterministic do not exhibit this sort of radical context dependence. A theory is simply deterministic or not. So, as it stands, the definition is hopeless: in one context a given W' and S' may be counterparts of W and S respectively; in another context, perhaps S' is a counterpart of S but W' is not a counterpart of W. Then according to our definition, W will be deterministic with respect to the first counterpart

[147] Here W' is physically possible with respect to W if they share the same laws of physics. We may (but need not) use Lewis's analysis of laws of nature: the laws of a world W are the axioms of an ideal scientific theory of W (Lewis 1983, pp. 365-68).

relation, and indeterministic with respect to the second. So the definition is unacceptable because it makes determinism into a context dependent feature of worlds.

If we are going to have an acceptable counterpart-theoretic analysis of determinism along the lines of Definition 0 then we need some way of giving a context independent specification of which counterpart relations count when we are checking to see whether or not a world is deterministic. Lewis offers such a specification (Lewis 1983, pp. 360-1). It depends on accepting that there are certain distinguished properties and relations that Lewis calls 'perfectly natural'; and then declaring that it is exactly those counterpart relations which preserve these perfectly natural properties and relations that matter when we are checking to see if a world is deterministic or not. Since the distinction between perfectly natural properties and relations and others is not context dependent, this gives us a context independent definition of determinism.

In general, Lewis's definition may be difficult to apply, since it requires that we know which properties and relations are perfectly natural. There is reason to hope, however, that the hole argument will present an exception to this rule, since, according to Lewis, spatiotemporal relations are clear examples of perfectly natural relations. Hence, only counterpart relations which preserve spatiotemporal relations will count when we check to see whether or not general relativity is deterministic. This is just what Butterfield needs to block the hole argument—spacetime points will be counterparts iff they share all of their geometric properties.

2. Method.

A few notes about methodology before I begin.

(i) Here I am interested in the problem of giving a counterpart-theoretic definition of determinism. I assume that we accept the rest of Lewis's account as true and focus my criticism on his definition of determinism. Perhaps his definition could be saved from my counterexamples by tinkering with the rest of his theory? Perhaps, but I doubt it. In any case, I will not discuss this question in this Appendix.

(ii) Some readers may be tempted to dismiss my counterexamples as involving bizarre worlds very far removed from our own. For such readers, I would like to point out two reasons for taking the examples seriously. (1) The laws of nature of my worlds are not so very different from our own. In fact, it is only their simplicity which makes them seem bizarre. (2) I follow Lewis in thinking that an analysis of a modal notion should hold at every possible world. Lewis 1986 is full of bizarre examples which he expects us to take seriously.

(iii) The examples work as follows. We have a proposed definition of determinism on the table. Then I bring out the hammer: a possible world which is deterministic according to the proposed definition but which I claim is indeterministic. In the first example, the hammer is easy to wield: I have a world which is clearly indeterministic. But in the second and third examples, I need a more delicate tool, since it will not be obvious that the worlds involved are indeterministic. I claim that the following is a necessary condition for a world W to be deterministic:

.

When (∗) fails, the statement is true in every context, and we seem to have a clear violation of our intuition that a world is deterministic iff every initial segment has only one physically possible future. In examples 2 and 3 below I will argue that (∗) is violated, and hence that the definitions on the table are unacceptable.

(iv) These arguments will depend on using counterpart relations under which things have counterparts other than themselves in their own worlds. For this reason, some readers may think that my arguments are essentially haecceitistic (i.e. that they appeal to some kind of Kripkean primitive ‘thisness’)—and thus that they should not be taken seriously, since haecceitism is incompatible with the spirit of Lewis’s approach to modality. My defense is that Lewis himself introduces samewordly counterparts as a way of doing justice to certain strong haecceitistic intuitions without admitting a notion of primitive thisness into his theory (Lewis 1986, pp. 227-35). My arguments are no more (or less) haecceitistic than Lewis's own.

3. Materials.

In order to state Lewis's definition, I must first remind you of some relatively specialized features of his account of modality.

Recall that for Lewis a property is just a set, the set of all its instances both actual and possible; and that any set of possibilia is a property in this sense (Lewis 1983, p. 343). Properties of this sort are *abundant*. They carve nature at the joints—but that is because they carve it everywhere, and in every way. Almost all of them are inconceivable for finite beings, and of the few that we could possibly grasp, only a minute proportion will ever play a role in our discourse. Grue and suchlike are utterly tame in the hierarchy of gerrymandered properties. Lewis claims that this is a useful explication of *one* of our intuitive notions of what a property is. But he notes that there are other intuitive notions which are just as important for philosophical purposes. In particular, we also have a notion of properties as being *sparse*. Lewis calls such properties *perfectly natural*. The perfectly natural properties carve nature at the joints, and *only* at the joints. Perfectly natural properties give us a way of making sense of resemblance: two things resemble each other if they share some perfectly natural properties. The abundant sort of property is no good for this sort of work: "Any two things share infinitely many properties, and fail to share infinitely many others. That is so whether the things are perfect duplicates or utterly dissimilar. Thus properties do nothing to capture the notion of resemblance" (p. 346). Lewis 1983 is a catalogue of the work that can be done by perfectly natural properties but not by ordinary properties. Lewis takes this to be a good reason to admit a distinction (which need not be sharp) between those properties which are perfectly natural, and those which are just plain old properties.[148] How can we know what the perfectly natural properties are? "Physics is relevant because it aspires to give an inventory of natural properties—not a complete inventory, perhaps, but a complete enough inventory to account for duplication among actual things. If physics succeeds in this, then duplication within our world amounts to sameness of physical description" (pp. 356-7). In our world there may be a relatively short list of perfectly natural properties—the properties which we use to classify elementary particles. In other worlds there may be very different sorts of perfectly natural properties (Lewis 1986, p. 60).

All of this is also true for relations: for Lewis a dyadic relation is just an arbitrary set of ordered pairs of possible things; this notion is of limited utility, so we introduce alongside it the notion of a perfectly natural relation. Paradigm examples of perfectly natural relations are the part-whole relation, and spatiotemporal relations.[149] These relations are perfectly natural for every possible world.[150]

Now we can define 'duplication': "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. 61). So (setting aside other perfectly natural relations, if there are any) if X is a duplicate of Y, and A is a duplicate of B, then the object composed of X and A is a duplicate of the object composed of Y an B iff the spatiotemporal relations between X an A are the same as those between Y an B.

---

[148] The distinction could have several origins: a primitive notion of a natural property, a primitive notion of resemblance, universals, or tropes. See pp. 347-8 of Lewis 1983 and pp. 63-9 of Lewis 1986.

[149] Indeed, Lewis speculates that these may be the *only* perfectly natural relations; Lewis 1986, p. 67.

[150] In fact, Lewis uses spatiotemporal relations to delimit possible worlds; see §1.6 of Lewis 1986.

Let W be a possible world. We can identify the initial segment of W up to time t with W_t, the set of all things which exist in W up until time t.[151] Now let $W'_{t'}$ be an initial segment of some other possible world W', an let $f:W_t \rightarrow W'_{t'}$ be a bijection. Then call f a duplication if f(x) is a duplicate of x for all x in W_t and f preserves all perfectly natural relations between components of compound objects.

The worlds that we will be interested in are ones in which very strong physicalist theses hold. So we may take W_t to include only spacetime points and temporal stages of particles—everything else will supervene on these. So for our purposes, a bijection $f:W_t \rightarrow W'_{t'}$ will be a duplication iff it preserves spatiotemporal relations and perfectly natural properties of spacetime points and temporal stages of particles.

Now we are ready for Lewis's definition: "I shall say that two worlds *diverge* iff they are not duplicates but they do have duplicate initial temporal segments" (Lewis 1983, p. 359); and, finally, determinism: "First, a system of laws of nature is Deterministic iff no two divergent worlds both conform perfectly to the laws of that system. Second, a world is Deterministic iff its laws comprise a Deterministic system" (p. 360). So we have

Definition 1: W is deterministic if, whenever W' is physically possible with respect to W and t, t', and $f:W_t \rightarrow W'_{t'}$ are such that f is a duplication, there is some duplication $g:W \rightarrow W'$.

4. Results.

Example 1: Collapse of columns.[152]

Imagine placing an object on top of a perfectly cylindrical metal column. If the object is not too heavy, the column will achieve an internal equilibrium. But if the weight of the object exceeds a certain critical weight then the column will collapse. The standard

[151] When Lewis speaks of a world W, or of one of its initial segments, he means a mereological sum rather than a set; Lewis 1986, p. 69 fn. 51. For my purposes, it is more convenient to work with sets. I don't think that anything hangs on this choice.
[152] This example is taken from pp. 215-6 of Wilson 1993, where it is used for a similar purpose.

engineering treatment of this phenomenon tells us to expect the column to collapse by buckling in a particular direction (see Bleich 1952, pp. 4-8). This is true even if the column is 'centrally loaded'—that is, even if we treat the object as a point placed on the axis of symmetry of the column. One may well doubt that this is a good description of what would happen to a centrally loaded column in our own world.[153] But I think that everyone will agree that there are possible worlds where perfectly symmetrical columns *do* collapse by buckling, even when centrally loaded. And I think that everyone should agree that such worlds should count as indeterministic, since the direction in which a given column will buckle is necessarily undetermined.[154]

But now consider a world W, in which centrally loaded columns collapse by buckling, which contains nothing but: (i) a large, homogenous, perfectly spherical planet; (ii) a relatively small perfectly cylindrical column resting on the planet so that its axis of symmetry is normal to the planet at the point of contact; (iii) and a cone, which is moving through space in such a way that at time t=0, its apex will impact the column at the exact center of its top surface, with sufficient force to cause the column to buckle. We also assume that W has a Newtonian spacetime structure; that the laws concerning motion are completely deterministic; and that the law governing column collapse fully determines the shape that the column assumes upon collapse.

It is easy to see that W counts as deterministic under Definition 1. Let t, W' and t' be such that there is a duplication $f:W_t \rightarrow W'_{t'}$.[155] Then f tells us that there are identical complete physical descriptions of W_t and $W'_{t'}$. So there must be identical complete physical descriptions of W and W': the laws have been chosen so that the only ambiguity in the evolution of the worlds is in the direction of collapse; but because of the symmetries of the worlds, it is always possible to give identical descriptions for them. So there is a duplication

[153] See Lecture 6 of Marsden 1981 for a treatment of this system as a (deterministic) Hamiltonian system..
[154] As discussed above, Brighouse 1995 declines to do so. See §1.
[155] If no such t, t', W' and f exist, then the definition will be satisfied vacuously, and W will count as deterministic. Here and elsewhere it is assumed that W' is physically possible with respect to W.

$g:W \to W'$, and W counts as deterministic under Definition 1. I take this as proof that there is something wrong with Lewis's definition.

The problem is clear: we want to be able to say that W and W' may differ in that the column may collapse in different directions in the two worlds; so we need some way of referring to directions in W and W'. The duplication $f:W_t \to W'_{t'}$ gives us this. Say that the column in W collapses into an elbow shape, and say that α is the particle on the outside of the elbow; if $t<0$ and a is the temporal stage of α at time t, then we can look at f(a) and the particle α' of which it is a stage in W', and see whether or not α' ends up on the outside of the elbow in W'. The fact that there exist f for which α' is not on the outside of the elbow licenses us to say that the column *could have* collapsed in another direction. Notice that given such an f, there is no duplication $g:W \to W'$ which extends f. I claim that it is this fact which is characteristic of the type of indeterminism which we see in this example. This leads us to

Definition 2: W is deterministic if, whenever W' is physically possible with respect to W, and t, 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.

Example 2: particle decay without spacetime points.

My second example involves a very simple kind of particle decay. There are some obvious ways in which particle decay can be an indeterministic process: (i) the time of decay could be undetermined; (ii) the products of decay could be undetermined; (iii) the directions that the products move in could be undetermined. Here I am not interested in any of these kinds of indeterminacy. To eliminate them, we postulate the following situation. A law which decrees that when the universe comes into existence, its only material contents will be particles of type α, which, after exactly 13 years, will each decay (this takes care of (i)). Every decay will produce two particles of type β (this takes care of (ii)). The initial velocities of the β particles will be equal in magnitude and separated by 80° in direction, and the center of mass of the two β particles will have the same velocity as the α particle which produced them; and our world W will initially contain a single particle of type α, and no spacetime points—spatiotemporal relations in W are properties of temporal stages of material objects which are not parasitic on relations among a substratum of spacetime points (these jointly take care of (iii)).

Let t, t', W' and f be such that $f{:}W_t \rightarrow W'_{t'}$ is a duplication. The case of interest to us is when t≤13. Label the particles $\alpha_1, \beta_1, \beta_2$ and $\alpha_1', \beta_1', \beta_2'$ in the obvious way. There is an extension of f to a duplication g_1 from W to W' which carries β_i onto β_i' (i=1,2). But there is also an extension of f to a duplication g_2 from W to W' which carries β_1 onto β_2' and β_2 onto β_1'.

According to Definition 2, W is a deterministic world: the fact that for every such f there is a g which extends f is sufficient—the fact that for each f there are actually two such g's is irrelevant. But is this right?

If W' is distinct from W, I don't think that we can make much of the ambiguity. But if W'=W, and $\beta_i = \beta_i'$, then g_2 seems to license us to say that β_1 could have been β_2. If we were somehow observing W, we could say that although as it turns out, β_1 is *this* particle, and β_2 *that* one, it could have been the other way around. I would say that W is indeterministic: before the decay, there are two possibilities for its future.

Perhaps this seems far-fetched. Consider Lewis's discussion of a samewordly counterpart relation in a slightly less extreme context (Lewis 1986, pp. 231-2).[156] "Here I

[156] I do not suggest that Lewis would endorse what I am about to say. On p. 232 of Lewis 1986, he emphasizes the distinction between a possibility and a possible world. He might well take a line advocated by Butterfield and Melia (private correspondence), and argue that although we have two possibilities in my case, we have only one possible world, and so no indeterminism. This amounts to insisting that a violation of (∗) by a counterfactual which is made true only by counterpart relations involving sameworldly counterparts does not count against the claim that a given world is deterministic.

Of course, we are owed an explanation of why it is only some possibilities that matter when it comes time to decide whether or not a world is deterministic. I expect that this would be a difficult task, even for a hard-core modal realist like Lewis. He, however, might be able to come up with such an explanation by leaning heavily on the (for him) am, there goes poor Fred; there but for the grace of God go I; how lucky I am to be me, not him." How can we make sense of this idea? Lewis claims that we should view Fred as being a counterpart of Lewis:

> Fred … is a possible way for a person to be. And in a sense he is even a possible way for me to be. He is my counterpart under an extraordinarily generous counterpart relation, one which demands nothing more of counterparts than that they be things of the same kind. Any property that one of my counterparts does have is a property that I might have; being Fred—being literally identical with him—is such a property; and so there is a sense in which I might have been him.

It is a generous counterpart relation indeed that makes Lewis the counterpart of an arbitrary Fred, simply by virtue of their shared humanity. Such a counterpart relation would be used only in fairly special conversational contexts—usually, the assertion that Lewis might have been Fred would be met with a blank stare. The assertion is true as true can be in certain contexts, but in most it is not. In that sense it is a tenuous possibility, and one that we should not be too concerned about.

Things are different with our world W. If you agree with Lewis's treatment of the Fred example, you will agree that there are some contexts in which it is true to say that β_1 could have been β_2. But now ask yourself in which contexts this would *not* be true. Counterpart relations follow qualitative similarity, and our example has been arranged so that β_1 and β_2 share *all* of their qualitative properties. So there can be no context in which it is false that β_1 could have been β_2. This is as robust a truth as any that there are. So W violates my condition (∗) of §2, and we must conclude that it is an indeterministic world.

The natural way to deal with example 2 would be to adopt

Definition 3: a world W is deterministic if, whenever W' is physically possible with respect to W, and t, t', W' and $f{:}W_t \rightarrow W'_{t'}$ are such that f is a duplication, then there is exactly one duplication $g{:}W \rightarrow W'$ which extends f.

---

metaphysically real distinction between possibilities and possible worlds. But for Butterfield, who wants to insist that his solution to the hole argument does not depend on an espousal of modal realism, the task will presumably be even more difficult—since to someone who does not accept modal realism, both possibilities and possible worlds will be mere fictions for doing modal semantics, so that the possibility/possible world distinction will not look very important.

Example 3: spherical particle decay with spacetime points.

In this example, W is a world with spacetime points and Newtonian spacetime structure. It initially contains a single α particle. The laws of nature decree that thirteen years later, at t=0, the α particle decays into continuum many β particles; arranged so that at time t, the β particles form a spherical shell of radius t; with each β particle moving away from the center of the sphere along its radius.

W counts as deterministic under our latest Definition. Let W', t, t', and $f:W_t \to W'_{t'}$ be such that f is a duplication. It is easy to convince yourself that every such f is extendible to a duplication $g:W \to W'$. To see that the extension is unique, we may work with the special case W'=W and t'=t, since there will be multiple extensions there iff there are multiple extensions in other cases (just compose with g or g^{-1}). Now, what does f do to the material contents of spacetime? Let c be some temporal stage of a particle, and let p be the spacetime point that it occupies in W. The distance between c and p is zero; distance is a spatiotemporal relation, and duplications respect such relations; so the distance between f(c) and f(p) is zero; therefore, f(c) occupies f(p). In particular, if a is the temporal stage of the α particle at time t and occupies spacetime point p; then f(a) must occupy f(p) at time t; the only occupied point at time t is p; so f fixes a and p. Since f preserves all spatiotemporal relations, it must restrict to a spacetime symmetry on the set S_t of spacetime points of W_t. Since W has Newtonian spacetime structure, this means that f will restrict to the same rotation about the center of mass of the system on each hyperplane of simultaneity on which it is defined.

But any duplication g which extends f must be a symmetry on S, the set of spacetime points of W. And, of course, there is only one such g. The action of g on material particles is determined by the requirement that it preserve the relation of occupation between temporal stages of particles and spacetime points. So W is deterministic according to Definition 3.

But isn't W indeterministic in the same way as the world of example 2? Fix β_1 and β_2, any two β particles. And let s be a symmetry of the spacetime structure of W that carries the worldline of β_1 onto the worldline of β_2. Define g':W→W as follows: g' restricted to S is the identity; for any particle stage c, g(c)=s(c). g' is not a duplication because it separates particle stages from the points they occupy. As in example 2, g' is a counterpart relation that tells us that β_1 *could have been* β_2 etc. g and g' license us to say that there are two possible futures for W_t: the one which actually obtains, and the one in which all the worldlines of the β particles are shifted by a rotation, while the spacetime points are fixed. And as in example 2, the two possibilities are qualitatively identical, so that there is no context in which it is false to say that the second future is a genuine possibility. So by principle (∗), W is indeterministic, and Definition 3 is inadequate.

5. Conclusions.

Definitions 1-3 are the only ways that I can think of to restrict the scope of Definition 0 to duplications. I think that the preceding examples show that none of these definitions is adequate as a counterpart-theoretic explication of determinism. So counterpart theorists interested in determinism have a choice: to forsake the form of Definition 0 with its close relationship to our intuitive notion of determinism; or to forsake their claim that it is duplications that matter when it comes to deciding whether or not a world is deterministic. Either way there is a lot of work to do—to defend a new form for a definition, or to attempt to single out another class of counterpart relations to do the work that duplications were supposed to do. Until this work is done, there is no counterpart-theoretic justification of Lockean substantivalism, and counterpart theory cannot rescue us from the 'threat' of the hole argument.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abraham, R. and J. E. Marsden (1978), *Foundations of Mechanics*. Reading, MA: The Benjamin/Cummings Publishing Company.

Aharonov, Y. (1986), "Non-Local Phenomena and the Aharonov-Bohm Effect," in R. Penrose and C. J. Isham (eds.), *Quantum Concepts in Space and Time*. Oxford: Clarendon Press, pp. 41-56.

Aharonov, Y. and D. Bohm (1959), "Significance of Electromagnetic Potentials in Quantum Theory," *Physical Review* 115: 485-91.

Alexander, H. G., (ed.) (1956), *The Leibniz-Clarke Correspondence: Together with Extracts From Newton's Principia and Opticks*. New York: Barnes and Noble.

Arnowitt, R., S. Deser, and C. W. Misner (1962), "The Dynamics of General Relativity," in L. Witten (ed.), *Gravitation: An Introduction to Current Research*. New York: Wiley, pp. 227-65.

Ashtekar, A. (1991), *Lectures on Non-Perturbative Gravity*. Singapore: World Scientific.

Ashtekar, A. (1994), "Overview and Outlook," in J. Ehlers and H. Friedrich (eds.), *Canonical Gravity: From Classical to Quantum*. New York: Springer-Verlag, pp. 327-67.

Ashtekar, A. (1995), "Mathematical Problems of Non-Perturbative Quantum General Relativity," in B. Julia and J. Zinn-Justin (eds.), *Gravitation and Quantizations*. Amsterdam: Elsevier, pp. 181-283.

Ashtekar, A., L. Bombelli, and O. Reula (1991), "The Covariant Phase Space of Asymptotically Flat Gravitational Fields," in M. Francaviglia (ed.), *Mechanics, Analysis and Geometry: 200 Years After Lagrange*. Amsterdam: North-Holland, pp. 417-50.

Ashtekar, A. and R. Geroch (1974), "Quantum Theory of Gravitation," *Reports on Progress in Physics* 37: 1211-56.

A Stachel, (eds.) (1991), *Conceptual Problems of Quantum Gravity*. Boston: Birkhäuser.

Ashtekar, A. and M. Stillerman (1986), "Geometric Quantization and Constrained Systems," *Journal of Mathematical Physics* 27: 1319-30.

Auyang, S. Y. (1995), *How Is Quantum Field Theory Possible*? Oxford: Oxford University Press.

Baez, J., (ed.) (1994), *Knots and Quantum Gravity*. Oxford: Clarendon Press.

Ballentine, L. (1996), “The Emergence of Classical Properties from Quantum Mechanics,” in R. Clifton (ed.), *Perspectives on Quantum Reality: Non-Relativistic, Relativistic, and Field-Theoretic*. Dordrecht: Kluwer, pp. 111-24.

Barbour, J. (1982), “Relational Concepts of Space and Time,” *British Journal for the Philosophy of Science* 33: 251-74.

Barbour, J. (1994a), “The Timelessness of Quantum Gravity: I. The Evidence from the Classical Theory,” *Classical and Quantum Gravity* 11: 2853-73.

Barbour, J. (1994b), “The Timelessness of Quantum Gravity: II. The Appearance of Dynamics in Static Configurations,” *Classical and Quantum Gravity* 11: 2875-97.

Barbour, J. (1995), “General Relativity as a Perfectly Machian Theory,” in J. Barbour and H. Pfister (eds.), *Mach's Principle: From Newton's Bucket to Quantum Gravity*. Boston: Birkhäuser, pp. 214-37.

Barbour, J. and B. Bertotti (1982), “Mach's Principle and the Structure of Dynamical Theories,” *Proceedings of the Royal Society of London* Series A 382: 295-306.

Bartels, A. (1994), “Traditional and Modern Essentialism and the Problem of Individuation of Spacetime Points,” Unpublished.

Batterman, R. (1995), “Theories Between Theories: Asymptotic Limiting Intertheoretic Relations,” *Synthese* 103: 171-205.

Bayle, P. (1965), *Historical and Critical Dictionary: Selections*. Indianapolis: Bobbs-Merrill.

Beig, R. (1994), “The Classical Theory of Canonical General Relativity,” in J. Ehlers and H. Friedrich (eds.), *Canonical Gravity: From Classical to Quantum*. New York: Springer-Verlag, pp. 59-80.

Bell, J. S. (1987), *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*. Cambridge: Cambridge University Press.

Belot, G. (1995a), “Determinism and Ontology,” *International Studies in the Philosophy of Science* 9: 85-101.

Belot, G. (1995b), “New Work for Counterpart Theorists: Determinism,” *British Journal for the Philosophy of Science* 46: 185-95.

Belot, G. and J. Earman (1996a), “Chaos Out of Order: Quantum Mechanics, the Correspondence Principle, and Chaos,” forthcoming in *Studies in History and Philosophy of Modern Physics* .

Belot, G. and J. Earman (1996b), “The Hole Argument, Gauge Groups, and the Concept of Observable,” Unpublished.

Bergmann, G. (1953), “The Logic of Quanta,” in H. Feigl and M. Brodbeck (eds.), *Readings in the Philosophy of Science*. New York: Appleton-Century-Crofts, pp. 475-508.

Bergson, H. (1910), *Time and Free Will: An Essay on the Immediate Data of Consciousness*. London: George Allen & Unwin Ltd.

Berry, M. (1994), "Asymptotics, Singularities and the Reduction of Theories," in D. Prawitz, B. Skyrms and D. Westerståhl (eds.), *Logic, Methodology, and Philosophy of Science IX*. Amsterdam: Elsevier, pp. 597-607.

Biagioli, M. (1990), "The Anthropology of Incommensurability," *Studies in History and Philosophy of Science* 21: 183-209.

Bleich, F. (1952), *Buckling Strength of Metal Structures*. New York: McGraw-Hill.

Boothby, W. M. (1986), *An Introduction to Differentiable Manifolds and Riemannian Geometry*. New York: Academic Press.

Boyd, R. (1984), "The Current Status of Scientific Realism," in J. Leplin (ed.), *Scientific Realism*. Berkeley: University of California Press, pp. 41-82.

Bricker, P. (1993), "The Fabric of Space: Intrinsic vs. Extrinsic Distance Relations," in P. A. French, T. E. Uehling and H. K. Wettstein (eds.), *Philosophy of Science*. Notre Dame: University of Notre Dame Press, pp. 271-95.

Brighouse, C. (1994), "Spacetime and Holes," in D. Hull, M. Forbes and R. M. Burian (eds.), *PSA 1994*, Vol. 1. East Lansing, MI: Philosophy of Science Association, pp. 117-25.

Brighouse, C. (1995), "Determinism and Modality," Unpublished.

Brown, H. R. and R. Harré, (eds.) (1988), *Philosophical Foundations of Quantum Field Theory*. Oxford: Clarendon Press.

Brown, J. R. (1994), *Smoke and Mirrors: How Science Reflects Reality*. New York: Routledge.

Buchwald, J. (1985), *From Maxwell to Microphysics: Aspects of Electromagnetic Theory in the Last Quarter of the Nineteenth Century*. Chicago: University of Chicago Press.

Buchwald, J. (1992), "Kinds and the Wave Theory of Light," *Studies in History and Philosophy of Science* 23: 39-74.

Butterfield, J. (1987), "Substantivalism and Determinism," *International Studies in the Philosophy of Science* 2: 10-32.

Butterfield, J. (1989a), "Albert Einstein Meets David Lewis," in M. Forbes and A. Fine (eds.), *PSA 1988*, Vol. Two. East Lansing MI: Philosophy of Science Association, pp. 65-81.

Butterfield, J. (1989b), "The Hole Truth," *British Journal for the Philosophy of Science* 40: 1-28.

Butterfield, J. (1993), "Interpretation and Identity in Quantum Theory," *Studies in History and Philosophy of Science* 24: 443-76.

Cao, T. Y. (1988), "Gauge Theory and the Geometrization of Fundamental Physics," in H. R. Brown and R. Harré (eds.), *Philosophical Foundations of Quantum Field Theory*. Oxford: Oxford University Press, pp. 117-33.

Carlip (1990), "Observables, Gauge Invariance, and Time in (2+1)-Dimensional Quantum Gravity," *Physical Review D* 42: 2647-54.

Carlip, S. (1991), "Time in 2+1 Dimensional Gravity," in R. Mann and P. Wesson (eds.), *Gravitation: A Banff Summer Institute*. Singapore: World Scientific, pp. 246-59.

Carlip, S. (1994), "Notes on the (2+1)-Dimensional Wheeler-DeWitt Equation," *Classical and Quantum Gravity* 11: 31-39.

Carlip, S. and J. E. Nelson (1995), "Comparative Quantizations of (2+1)-Dimensional Gravity," *Physical Review D* 51: 5643-53.

Carnap, R. (1958), *Introduction to Symbolic Logic*. New York: Dover.

Carrol, J. W. (1994), *Laws of Nature*. Cambridge: Cambridge University Press.

Cartwright, N. (1983), *How the Laws of Physics Lie*. Oxford: Oxford University Press.

Cassa, A. (1996), "Reconstructing Space-Time from World-Lines of Charged Particles," *Journal of Geometry and Physics* 19: 77-89.

Cassirer, E. (1956), *Determinism and Indeterminism in Modern Physics: Historical and Systematic Studies of the Problem of Causality*. New Haven: Yale University Press.

Choquet-Bruhat, Y., C. DeWitt-Morrette, and M. Dillard-Bleick (1982), *Analysis, Manifolds, and Physics*. Amsterdam: North-Holland.

Christian, J. (1994), "On Definite Events in a Generally Covariant Quantum World," Unpublished.

Connes, A. and C. Rovelli (1994), "von Neumann Algebra Automorphisms and Time-Thermodynamics Relation in Generally Covariant Quantum Theories," *Classical and Quantum Gravity* 11: 2899-2917.

Cosgrove, R. (1996), "Consistent Evolution and Different Time Slicings in Quantum Gravity," *Classical and Quantum Gravity* 13: 891-919.

Cushing, J. (1990), "Foundational Problems in and Methodological Lessons from Quantum Field Theory," in R. Harré and H. Brown (eds.), *Philosophical Foundations of Quantum Field Theory*. Oxford: Clarendon Press, pp. 25-39.

Cushing, J. (1996), "What Measurement Problem?," in R. Clifton (ed.), *Perspectives on Quantum Reality: Non-Relativistic, Relativistic, and Field-Theoretic*. Dordrecht: Kluwer, pp. 167-81.

Davidson, D. (1980), *Essays on Actions and Events*. Oxford: Clarendon Press.

De León, M. and P. R. Rodrigues (1985), *Generalized Classical Mechanics: A Geometrical Approach of Lagrangian and Hamiltonian Formalisms Involving Higher Order Derivatives*. Amsterdam: North-Holland.

Descartes, R. (1967), *The Philosophical Works of Descartes*. Cambridge: Cambridge University Press.

Dirac, P. A. M. (1964), *Lectures on Quantum Mechanics*. New York: Yeshiva University.

DiSalle, R. (1994), "On Dynamics, Indiscernability, and Spacetime Ontology," *British Journal for the Philosophy of Science* 45: 265-87.

DiSalle, R. (1995), "Spacetime Theory as Physical Geometry," *Erkenntnis* 42: 317-37.

Dutens, L., (ed.) (1768), *G.G. Leibnitii Opera Omnia*. Geneva: Fratres de Tournes.

Duval, C. and H. P. Künzle (1984), "Minimal Gravitational Coupling in the Newtonian Theory and the Covariant Schrödinger Equation," *General Relativity and Gravitation* 16: 333-47.

Earman, J. (1970a), "Space-Time, or How to Solve Philosophical Problems and Dissolve Philosophical Muddles Without Really Trying," *Journal of Philosophy* 67: 259-77.

Earman, J. (1970b), "Who's Afraid of Absolute Space," *Australasian Journal of Philosophy* 48: 287-317.

Earman, J. (1986), *A Primer on Determinism*. Dordrecht: Reidel.

Earman, J. (1989), *World Enough and Space-Time*. Cambridge, MA: MIT Press.

Earman, J. (1995), *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes*. Oxford: Oxford University Press.

Earman, J. and M. Friedman (1973), "The Meaning and Status of Newton's Law of Inertia and the Nature of Gravitational Forces," *Philosophy of Science* 40: 329-59.

Earman, J. and J. Norton (1987), "What Price Spacetime Substantivalism? The Hole Story," *British Journal for the Philosophy of Science* 38: 515-25.

Einstein, A. (1952), "The Foundation of the General Theory of Relativity," in W. Perrett and G. B. Jeffery (eds.), *The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity*. New York: Dover, pp. 109-64.

Emch, G. G., H. Narnhofer, W. Thirring, and G. L. Sewell (1994), "Anosov Actions on Noncommutative Algebras," *Journal of Mathematical Physics* 35: 5582-99.

Feigl, H. (1953), "Notes on Causality," in H. Feigl and M. Brodbeck (eds.), *Readings in the Philosophy of Science*. New York: Appleton-Century-Crofts, pp. 408-18.

Field, H. (1980), *Science Without Numbers*. Princeton: Princeton University Press.

Field, H. (1985), "Can We Dispense With Space-Time," in P. Asquith and P. Kitcher (eds.), *PSA 1984*, Vol. Two. East Lansing, MI: Philosophy of Science Association, pp. 33-90.

Fine, A. (1984a), "And Not Anti-Realism Either," *Noûs* 18: 51-65.

Fine, A. (1984b), "The Natural Ontological Attitude," in J. Leplin (ed.), *Scientific Realism.* Berkeley: University of California Press, pp. 83-107.

Fine, A. (1996), "Science Made Up: Constructivist Sociology of Scientific Knowledge," in P. Galison and D. J. Stump (eds.), *The Disunity of Science: Boundaries, Contexts, and Power*. Stanford: Stanford University Press, pp. 231-54.

Fock, V. (1964), *The Theory of Space, Time, and Gravitation*. New York: Pergamon Press.

Frank, P. (1957), *Philosophy of Science: The Link Between Philosophy and Science.* Englewood Cliffs, NJ: Prentice-Hall.

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

Friedman, M. (1993), "Remarks on the History of Science and the History of Philosophy," in P. Horwich (ed.), *World Changes: Thomas Kuhn and the Nature of Science*. Cambridge, MA: MIT Press, pp. 37-54.

Galison, P. (1995), "Context and Constraints," in J. Buchwald (ed.), *Scientific Practice: Theories and Stories of Doing Physics*. Chicago: University of Chicago Press, pp. 13-41.

Glymour, C. (1977), "The Epistemology of Geometry," *Noûs* 11: 227-251.

Göckeler, M. and T. Schücker (1987), *Differential Geometry, Gauge Theories, and Gravity*. Cambridge: Cambridge University Press.

Grant, E., (ed.) (1974), *A Source Book in Medieval Science*. Cambridge, MA: Harvard University Press.

Grant, E. (1979), "The Condemnation of 1277, God's Absolute Power, and Physical Thought in the Late Middle Ages," *Viator* 10: 211-44.

Grant, E. (1981), *Much Ado About Nothing: Theories of Space And Vacuum from the Middle Ages to the Scientific Revolution.* Cambridge: Cambridge University Press.

Grünbaum, A. (1973), *Philosophical Problems of Space and Time*. Dordrecht: Reidel.

Grünbaum, A. (1977), "Absolute and Relational Theories of Space and Space-Time," in J. Earman, C. Glymour and J. Stachel (eds.), *Foundations of Space-Time Theories*. Minneapolis: University of Minnesota Press, pp. 303-73.

Guillemin, V. and S. Sternberg (1984), *Symplectic Techniques in Physics*. Cambridge: Cambridge University Press.

Haag, R. (1992), *Local Quantum Physics: Fields, Particles, Algebras*. New York: Springer-Verlag.

Hacking, I. (1983), *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press.

Hacking, I. (1990), *The Taming of Chance*. Cambridge: Cambridge University Press.

Hacking, I. (1996), "The Disunities of the Sciences," in P. Galison and D. Stump (eds.), *The Disunity of Science: Boundaries, Contexts, and Power*. Stanford: Stanford University Press, pp. 37-74.

Hahn, R. (1967), *Laplace as a Newtonian Scientist: A Paper delivered at a Seminar on The Newtonian Influence held at the Clark Library, 8 April, 1967*. Los Angeles: UCLA Press.

Hajicek, P. (1991), "Comment on "Time in Quantum Gravity: An Hypothesis"," *Physical Review D* 44: 1337-38.

Hajicek, P. (1996), "Time Evolution and Observables in Constrained Systems," *Classical and Quantum Gravity* 13: 1353-75.

Hajicek, P. and C. Isham (1996a), "Perennials and the Group-Theoretic Quantization of a Parameterized Scalar Field on a Curved Background," *Journal of Mathematical Physics* 37: 3522-38.

Hajicek, P. and C. Isham (1996b), "The Symplectic geometry of a Parameterized Scalar Field on a Curved Background," *Journal of Mathematical Physics* 37: 3505-21.

Hall, A. R. (1980), *Philosophers at War*. Cambridge: Cambridge University Press.

Hall, A. R. and M. B. Hall, (eds.) (1962), *Unpublished Scientific Papers of Isaac Newton*. Cambridge: Cambridge University Press.

Hartle, J. (1996), "Time and Time Functions in Parameterized Non-Relativistic Quantum Mechanics," *Classical and Quantum Gravity* 13: 361-75.

Hawking, S. and G. F. R. Ellis (1973), *The Large Scale Structure of Space-Time*. Cambridge: Cambridge University Press.

Healey, R. (1995), "Substance, Modality and Spacetime," *Erkenntnis* 42: 287-316.

Henneaux, M. and C. Teitelboim (1992), *Quantization of Gauge Systems*. Princeton: Princeton University Press.

Hesse, M. (1961), *Forces and Fields: The Concept of Action at a Distance in the History of Physics*. London: Thomas Nelson and Sons.

Hoefer, C. (1996), "The Metaphysics of Space-Time Substantivalism," *Journal of Philosophy* XCIII: 5-27.

Hoefer, C., and N. Cartwright (1993), "Substantivalism and the Hole Argument," in J. Earman, A. Janis, G. Massey and N. Rescher (eds.), *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolph Grünbaum*. Pittsburgh: University of Pittsburgh Press, pp. 23-43.

Horvathy, P. A. (1980), "Quantisation in Multiply Connected Spaces," *Physics Letters* 76A: 11-14.

Horwich, P. (1978), "On the Existence of Time, Space, and Space-Time," *Noûs* 12: 397-419.

Howard, D. (1992), "Einstein and *Eindeutigkeit*: A Neglected Theme in the Philosophical Background to General Relativity," in J. Eisenstaedt and A. Kox (eds.), *Studies in the History of General Relativity*. Boston: Birkhäuser, pp. 154-243.

Hume, D. (1978), *A Treatise of Human Nature*. Oxford: Oxford University Press.

Isenberg, J. and V. Moncrief (1996), "A Set of Nonconstant Mean Curvature Solutions of the Einstein Constraint Equations on Closed Manifolds," *Classical and Quantum Gravity* 13: 1819-47.

Isham, C. (1991), "Conceptual and Geometrical Problems in Quantum Gravity," in H. Mitter and Gausterer (eds.), *Recent Aspects of Quantum Fields*. New York: Springer-Verlag, pp. 123-229.

Isham, C. (1993), "Canonical Quantum Gravity and the Problem of Time," in L. A. Ibort and M. A. Rodriguez (eds.), *Integrable Systems, Quantum Groups, and Quantum Field Theories*. Dordrecht: Kluwer, pp. 157-288.

Isham, C. (1994), "Prima Facie Questions in Quantum Gravity," in J. Ehlers and H. Friedrich (eds.), *Canonical Gravity: From Classical to Quantum*. New York: Springer-Verlag, pp. 1-21.

Jacobson, T. (1991), "Black-Hole Thermodynamics and the Space-Time Discontinuum," in A. Ashtekar and J. Stachel (eds.), *Conceptual Problems of Quantum Gravity*. Boston: Birkhäuser, pp. 597-600.

James, W. (1896), *The Will to Believe and Other Essays in Popular Philosophy*. New York: Longmans Green and Co.

Kennedy, J. B. (1993), "Interpreting Gauge Symmetries: a Proposed Solution to the Problem of Gauge Invariance," Unpublished.

Koyré, A. and I. B. Cohen (1962), "Newton and the Leibniz-Clarke Correspondence," *Archives Internationales d'Histoire des Sciences* 15: 63-126.

Kripke, S. (1980), *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Kuchar, K. (1980), "Gravitation, Geometry, and Nonrelativistic Quantum Theory," *Physical Review D* 22: 1285-99.

Kuchar, K. (1981), "Canonical Methods of Quantization," in C. J. Isham, R. Penrose and D. W. Sciama (eds.), *Quantum Gravity 2: A Second Oxford Symposium*. Oxford: Clarendon Press, pp. 329-76.

Kuchar, K. (1988), "Canonical Quantization of Generally Covariant Systems," in B. Iyer, A. Kembhavi, J. Narlikar and C. Vishveshwara (eds.), *Highlights in Gravitation and Cosmology*. Cambridge: Cambridge University Press, pp. 93-120.

Kuchar, K. (1991), "The Problem of Time in Canonical Quantization of Relativistic Systems," in A. Ashtekar and J. Stachel (eds.), *Conceptual Problems of Quantum Gravity*. Boston: Birkhäuser, pp. 141-71.

Kuchar, K. (1992), "Time and Interpretations of Quantum Gravity," in G. Kunsatter, D. Vincent and J. Williams (eds.), *Proceedings of the 4th Canadian Conference on General Relativity and Astrophysics*. Singapore: World Scientific, pp. 211-314.

Kuchar, K. (1993a), "Canonical Quantum Gravity," in R. J. Gleiser, C. N. Kozameh and O. M. Moreschi (eds.), *General Relativity and Gravitation 1992: Proceedings of the Thirteenth International Conference on General Relativity and Gravitation held at Cordoba, Argentina, 28 June-4 July 1992*. Philadelphia: Institute of Physics Publishing, pp. 119-50.

Kuchar, K. (1993b), "Matter Time in canonical Quantum Gravity," in B. Hu, M. Ryan and C. Vishveshwara (eds.), *Directions in General Relativity*. Cambridge: Cambridge University Press, pp. 201-21.

Kuhn, T. (1970), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kuhn, T. (1991), "The Road Since Structure," in A. Fine, M. Forbes and L. Wessels (eds.), *PSA 1990*, Vol. 2. East Lansing, MI: Philosophy of Science Association, pp. 3-13.

Laplace, P. S. (1951), *A Philosophical Essay on Probabilities*. New York: Dover.

Laudan, L. (1984), "A Confutation of Convergent Realism," in L. Leplin (ed.), *Scientific Realism*. Berkeley: University of California Press, pp. 218-49.

Leeds, S. (1995), "Holes and Determinism: Another Look," *Philosophy of Science* 62: 425-437.

Leibniz, G. W. (1981), *New Essays Concerning Human Understanding*. Cambridge: Cambridge University Press.

Leplin, J. (1984), "Truth and Scientific Progress," in J. Leplin (ed.), *Scientific Realism*. Berkeley: University of California Press, pp. 193-217.

Lewis, D. (1973), *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, D. (1983), "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 61: 343-77.

Lewis, D. (1986), *On the Plurality of Worlds*. Oxford: Basil Blackwell.

Liu, C. (1996), "Realism and Spacetime: Of Arguments Against Metaphysical Realism and Manifold Realism," forthcoming in *Philosophia Naturalis*.

Locke, J. (1975), *An Essay Concerning Human Understanding*. Oxford: Clarendon Press.

Loll, R. (1994), "The Loop Formulation of Gauge Theory and Gravity," in J. C. Baez (ed.), *Knots and Quantum Gravity*. Oxford: Clarendon Press, pp. 1-19.

Lynden-Bell, D. (1995), "A Relative Newtonian Mechanics," in J. Barbour and H. Pfister (eds.), *Mach's Principle: From Newton's Bucket to Quantum Gravity*. Boston: Birkhäuser, pp. 172-78.

Madore, J. (1996), "Linear Connections on Fuzzy Manifolds," *Classical and Quantum Gravity* 13: 2109-19.

Maidens, A. (1993), *The Hole Argument: Substantivalism and Determinism in General Relativity*. Unpublished Ph.D. thesis, University of Cambridge.

Malament, D. (1976), "Review of *Space, Time, and Spacetime*," *Journal of Philosophy* LXXIII: 306-23.

Malament, D. (1982), "Review of *Science Without Numbers*," *Journal of Philosophy* LXXIX: 523-34.

Malament, D. (1986), "Why Space Must Be Euclidean," in R. G. Colodny (ed.), *From Quarks to Quasars*. Pittsburgh: University of Pittsburgh Press, pp. 181-203.

Malament, D. (1995), "Is Newtonian Cosmology Really Inconsistent?," *Philosophy of Science* 62: 489-510.

Manders, K. (1982), "On the Space-Time Ontology of Physical Theories," *Philosophy of Science* 49: 575-90.

Marsden, J. E. (1981), *Lectures on Geometric methods in Mathematical Physics*. Philadelphia: SIAM.

Marsden, J. E. (1992), *Lectures on Mechanics*. Cambridge: Cambridge University Press.

Matschull, H.-J. (1995), "Three-Dimensional Canonical Quantum Gravity," *Classical and Quantum Gravity* 12: 2621-2703.

Maudlin, T. (1989), "The Essence of Space-Time," in M. Forbes and A. Fine (eds.), *PSA 1988*, Vol. Two. East Lansing MI: Philosophy of Science Association, pp. 82-91.

Maudlin, T. (1990), "Substances and Space-Time: What Aristotle Would Have Said to Einstein," *Studies in History and Philosophy of Science* 21: 531-61.

Maudlin, T. (1993), "Buckets of Water and Waves of Space: Why Spacetime Probably is a Substance," *Philosophy of Science* 60: 183-203.

Maudlin, T. (1994), *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*. Oxford: Blackwell.

M 1978), "Newton on Place, Time, and God: An Unpublished Source," *British Journal for the History of Science* 11: 114-29.

McGuire, J. E. (1982), "Space, Infinity, and Indivisibility: Newton on the Creation of Matter," in Z. Bechler (ed.), *Contemporary Newtonian Research*. Dordrecht: D. Reidel, pp. 145-90.

Minkowski, H. (1952), "Space and Time," in W. Perrett and G. B. Jeffery (eds.), *The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity*. New York: Dover, pp. 73-91.

Montague, R. (1974), *Formal Philosophy*. New Haven: Yale University Press.

Mundy, B. (1983), "Relational Theories of Euclidean Space and Minkowski Spacetime," *Philosophy of Science* 50: 205-26.

Mundy, R. (1992), "Space-Time and Isomorphism," in D. Hull, M. Forbes and K. Okruhlik (eds.), *PSA 1992*, Vol. 1. East Lansing, Michigan: Philosophy of Science Association, pp. 515-27.

Nagel, E. (1961), *The Structure of Science: Problems in the Logic of Scientific Explanation*. London: Routledge & Kegan Paul.

Nash, C. and S. Sen (1983), *Topology and Geometry for Physicists*. New York: Academic Press.

Nerlich, G. (1979), "What Can Geometry Explain?," *British Journal for the Philosophy of Science* 33: 361-82.

Nerlich, G. (1991), "How Euclidean Geometry has Misled Metaphysics," *Journal of Philosophy* 88: 169-89.

Nerlich, G. (1993), "Holes in the Hole Argument," in G. Nerlich (ed.), *What Spacetime Explains: Metaphysical Essays on Space and Time*. Cambridge: Cambridge University Press, pp. 206-23.

Nersessian, N. (1984), *Faraday to Einstein: Constructing Meaning in Scientific Theories*. Dordrecht: Martinus Nijoff Publishers.

Newton, I. (1952), *Opticks*. New York: Dover.

Norton, J. (1984), "How Einstein Found His Field Equations: 1912-1915," *Historical Studies in the Physical Sciences* 14: 253-316.

Norton, J. (1989), "The Hole Argument," in M. Forbes and A. Fine (eds.), *PSA 1988*, Vol. Two. East Lansing MI: Philosophy of Science Association, pp. 56-64.

Norton, J. (1992), "Philosophy of Space and Time," in M. H. Salmon (ed.), *Introduction to the Philosophy of Science: A Text by the Members of the Department of the History and Philosophy of Science of the University of Pittsburgh*. Englewood Cliffs, NJ: Prentice Hall, pp. 179-231.

Norton, J. (1993), "General Covariance and the Foundations of General Relativity," *Reports on Progress in Physics* 56: 791-858.

Norton, J. (1995), "The Force of Newtonian Cosmology: Acceleration is Relative," *Philosophy of Science* 62: 511-22.

Page, D. N. (1995), "The Hawking-Hartle Proposal for the Quantum State of the Universe," in J. Leplin (ed.), *The Creation of Ideas in Physics: Studies for a Methodology of Theory Construction*. Dordrecht: Reidel, pp. 181-88.

Peirce, C. S. (1923), *Chance, Love, and Logic*. New York: Harcourt, Brace & Company.

Penrose, R. (1996), "On Gravity's Role in Quantum State Reduction," *General Relativity and Gravitation* 28: 581-600.

Puta, M. (1993), *Hamiltonian Mechanical Systems and Geometric Quantization.* Dordrecht: Kluwer.

Redhead, M. (1987), *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford: Clarendon Press.

Redhead, M. (1990), “A Philosopher Looks at Quantum Field Theory,” in H. R. Brown and R. Harré (eds.), *Philosophical Foundations of Quantum Field Theory*. Oxford: Clarendon Press, pp. 9-23.

Reichenbach, H. (1959), “The Theory of Motion According to Newton, Leibniz, and Huyghens,” in M. Reichenbach (ed.), *Modern Philosophy of Science: Selected Essays*. London: Routledge & Kegan Paul, pp. 46-66.

Rendall, A. D. (1994), “Canonical Gravity: from Clasical to Quantum,” in J. Ehlers and H. Friedrich (eds.), New York: Springer-Verlag, pp. 319-26.

Rohrlich, F. (1988), “Pluralistic Ontology and Theory Reduction in the Physical Sciences,” *British Journal for the Philosophy of Science* 39: 295-312.

Rohrlich, F. (1996), “Scientific Realism: A Challenge to Physicists,” *Foundations of Physics* 26: 443-51.

Rovelli, C. (1991a), “Is There Incompatibility Between the Way Time is Treated in General Relativity and in Standard Quantum Mechanics?,” in A. Ashtekar and J. Stachel (eds.), *Conceptual Problems of Quantum Gravity*. Boston: Birkhäuser, pp. 126-40.

Rovelli, C. (1991b), “Quantum Evolving Constants. Reply to “Comment on ‘Time in Quantum Gravity: An Hypothesis’”,” *Physical Review D* 44: 1339-41.

Rovelli, C. (1991c), “Quantum Reference Systems,” *Classical and Quantum Gravity* 8: 317-31.

Rovelli, C. (1991d), “Time in Quantum Gravity: An Hypothesis,” *Physical Review D* 43: 442-56.

Rovelli, C. (1991e), “What is Observable in Classical and Quantum Gravity,” *Classical and Quantum Gravity* 8: 297-316.

Rovelli, C. (1993), “Statistical Mechanics of Gravity and the Thermodynamical Origin of Time,” *Classical and Quantum Gravity* 10: 1549-66.

Rovelli, C. (1994), “Half Way Through the Woods: Contemporary Research on Space and Time,” Unpublished.

Rynasiewicz, R. (1992), “Rings, Holes and Substantivalism: on the Program of Leibniz Algebras,” *Philosophy of Science* 59: 572-89.

Rynasiewicz, R. (1994), “The Lessons of the Hole Argument,” *British Journal for the Philosophy of Science* 45: 407-436.

Rynasiewicz, R. (1996a), “Absolute versus Relational Space-Time: An Outmoded Debate?,” *Journal of Philosophy* XCIII: 279-306.

Rynasiewicz, R. (1996b), "Is There a Syntactic Solution to the Hole Problem?," forthcoming in *PSA 1996*.

Salmon, W. (1984), *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Saunders, S. (1988), "The Algebraic Approach to Quantum Field Theory," in H. R. Brown and R. Harré (eds.), *Philosophical Foundations of Quantum Filed Theory*. Oxford: Clarendon Press, pp. 149-86.

Schmid, R. (1987), *Infinite Dimensional Hamiltonian Systems*. Naples: Bibliopolis.

Scriven, M. (1957), "The Present Status of Determinism in Physics," *Journal of Philosophy* 54: 727-41.

Sklar, L. (1974), *Space, Time, and Spacetime*. Berkeley: University of California Press.

Sklar, L. (1976), "Inertia, Gravitation and Metaphysics," *Philosophy of Science* 43: 1-23.

Smart, J. J. C. (1968), *Between Philosophy and Science*. New York: Random House.

Smolin, L. (1991), "Space and Time in the Quantum Universe," in A. Ashtekar and J. Stachel (eds.), *Conceptual Problems of Quantum Gravity*. Boston: Birkhäuser, pp. 228-91.

Sniatycki, J. (1980), *Geometric Quantization and Quantum Mechanics*. New York: Springer-Verlag.

Sorkin, R. D. (1995), "A Specimen of Theory Construction from Quantum Gravity," in J. Leplin (ed.), *The Creation of Ideas in Physics: Studies for a Methodology of Theory Construction*. Dordrecht: Reidel, pp. 167-80.

Stachel, J. (1989), "Einstein's Search for General Covariance, 1912-1915," in D. Howard and J. Stachel (eds.), *Einstein and the History of General Relativity*. Boston: Birkhäuser, pp. 63-100.

Stachel, J. (1993), "The Meaning of General Covariance," in J. Earman, A. Janis, G. Massey and N. Rescher (eds.), *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolph Grünbaum*. Pittsburgh: University of Pittsburgh Press, pp. 129-60.

Stein, H. (1967), "Newtonian Space-Time," *Texas Quarterly* 10: 174-200.

Stein, H. (1970), "On the Notion of Field in Newton, Maxwell, and Beyond," in R. Stuewer (ed.), *Historical and Philosophical Perspectives of Science*. Minneapolis: University of Minnesota Press, pp. 264-86.

Stein, H. (1977), "Some Philosophical Prehistory of General Relativity," in J. Earman, C. Glymour and J. Stachel (eds.), *Foundations of Space-Time Theories*. Minneapolis: University of Minnesota Press, pp. 3-49.

Stein, H. (1990), "On Locke, 'the Great Huygenius, and the Incomparable Mr. Newton'," in P. Bricker and R.I.G. Hughes (eds.), *Philosophical Perspectives on Newtonian Science*. Cambridge, MA: MIT Press, 17-47.

Stein, H. (1993), "On Philosophy and Natural Philosophy in the Seventeenth Century," in P. A. French, T. E. Uehling and H. K. Wettstein (eds.), *Philosophy of Science*. Notre Dame: University of Notre Dame Press, pp. 177-201.

Teller, P. (1987), "Space-Time as a Physical Quantity," in P. Achinstein and R. Kagon (eds.), *Kelvin's Baltimore Lectures and Modern Theoretical Physics*. Cambridge: MIT Press, pp. 425-48.

Teller, P. (1991), "Substance, Relations, and Arguments About the Nature of Space-Time," *The Philosophical Review* C: 362-97.

Teller, P. (1995), *An Interpretative Introduction to Quantum Field Theory*. Princeton: Princeton University Press.

Torre, C. G. (1993), "Gravitational Observables and Local Symmetries," *Physical Review D* 48: R2373-76.

Unruh, W. (1991), "No Time and Quantum Gravity," in R. Mann and P. Wesson (eds.), *Gravitation: A Banff Summer Institute*. Singapore: World Scientific, pp. 260-75.

Unruh, W. (1995), "Time, Gravity, and Quantum Mechanics," in S. F. Savitt (ed.), *Time's Arrows Today: Recent Physical and Philosophical Work on the Direction of Time*. Cambridge: Cambridge University Press, pp. 23-65.

Unruh, W. and R. Wald (1984), "What Happens When an Accelerating Observer Detects a Rindler Particle," *Physical Review D* 29: 1047-56.

Unruh, W. and R. Wald (1989), "Time and the Interpretation of Canonical Quantum Gravity," *Physical Review D* 1989: 2598-2614.

van Fraassen, B. (1980), *The Scientific Image*. Oxford: Clarendon Press.

van Fraassen, B. (1995), "'World' is Not a Count Noun," *Noûs* XXIX: 139-157.

Wald, R. (1984), *General Relativity*. Chicago: University of Chicago Press.

Wald, R. (1994), *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. Chicago: University of Chicago Press.

Weingard, R. (1996), "Exotic (Quixotic?) Applications of Bohm Theory," in R. Clifton (ed.), *Perspectives on Quantum Reality: Non-Relativistic, Relativistic, and Field-Theoretic*. Dordrecht: Kluwer, pp. 195-210.

Wilson, M. (1989), "Critical Notice: John Earman's *A Primer on Determinism*," *Philosophy of Science* 56: 502-32.

Wilson, M. (1993), "There's a Hole and a Bucket, Dear Leibniz," in P. A. French, T. E. Uehling and H. K. Wettstein (eds.), *Philosophy of Science*. Notre Dame: University of Notre Dame Press, pp. 202-41.

Woodhouse, N. (1980), *Geometric Quantization*. Oxford: Clarendon Press.