Pre-Socratic Quantum Gravity*

Gordon Belot   John Earman

1 Introduction

Physicists who work on canonical quantum gravity will sometimes remark that the general covariance of general relativity is responsible for many of the thorniest technical and conceptual problems in their field.[1] In particular, it is sometimes alleged that one can trace to this single source a variety of deep puzzles about the nature of time in quantum gravity, deep disagreements surrounding the notion of 'observable' in classical and quantum gravity, and deep questions about the nature of the existence of spacetime in general relativity.

Philosophers who think about these things are sometimes skeptical about such claims. We have all learned that Kretschmann was quite correct to urge against Einstein that the "General Theory of Relativity" was no such thing, since *any* theory could be cast in a generally covariant form, and hence the general covariance of general relativity could not have any physical content, let alone bear the kind of weight that Einstein expected it to.[2] Friedman's assessment is widely accepted: "As Kretschmann first pointed out in 1917, the principle of general covariance has no physical content whatever: it specifies no particular physical theory; rather, it merely expresses our commitment to a certain style of formulating physical theories" (1984, p. 44). Such considerations suggest that general covariance, as a technically crucial but physically contentless feature of general relativity, simply cannot be the source of any significant conceptual or physical problems.[3]

Physicists are, of course, conscious of the weight of Kretschmann's points against Einstein. Yet they are considerably more ambivalent than their philosophical colleagues. Consider Kuchař's conclusion at the end of a discussion of this topic:

---

*We would like to thank Karel Kuchař and Carlo Rovelli for invaluable tutelage, and Craig Callender, Nick Huggett, and Steve Weinstein for helpful comments.

[1]Here and throughout we restrict our attention to the canonical approach to quantum gravity.

[2]See Norton 1993 for a discussion of this and other episodes in the long debate over general covariance.

[3]See Belot and Earman 1999 for a discussion of some related contrasts between the pessimistic attitudes of (many) philosophers and the optimistic attitudes of (some) physicists with respect to the relevance of interpretative work on general relativity to ongoing research on quantum gravity.

> the Einstein-Kretschmann discussion is clearly relevant for the canonical quantization of covariant theories, but, as so many times before in the ancient controversy between the relative and the absolute, it is difficult to decide which of these two alternative standpoints is correct and fruitful. This leaves the canonical quantization of covariant systems uncomfortably suspended between the relative and the absolute. (1988, p. 118)

It becomes clear in the course of Kuchař's discussion that he takes the physical content of the general covariance of general relativity to reside not in the fact that that theory, like every other, *can* be given a generally covariant formulation, but in the fact that it *ought* to be so formulated (see, e.g., pp. 95-6). The idea is that one does some sort of violence to the physical content of general relativity if one breaks its general covariance by introducing preferred coordinates, slicings, or other geometrical structure, in a way in which one does not when one moves from a generally covariant formulation of Newtonian mechanics or special relativity to the standard formulations in which inertial coordinates play a special role.

Central to this way of thinking about general covariance is the idea that misjudging the physical content of a given theory can lead one astray in attempts to construct new theories—since, e.g., empirically equivalent formulations of a given theory may well lead to inequivalent quantum theories, it is important to begin with the *correct* formulation. This link between content and method is the source of the sentiment, widespread among physicists working on canonical quantum gravity, that there is a tight connection between the interpretative problems of general relativity and the technical and conceptual problems of quantum gravity.

Our goal in this paper is to explicate this connection for a philosophical audience, and to evaluate some of the interpretative arguments which have been adduced in favor of various attempts to formulate quantum theories of gravity. We organize our discussion around the question of the extent to which the general covariance of general relativity can (or should) be understood by analogy to the gauge invariance of theories like classical electromagnetism, and the related questions of the nature of observables in classical and quantum gravity, and the existence of time and change in the quantum theory.

We provide neither a comprehensive introduction to the formalism of quantum gravity, nor a survey of its interpretative problems (readers interested in the latter should turn to the canonical survey articles: Isham 1991, Isham 1993, and Kuchař 1992). We do, however, want the paper to be both accessible to readers who are unfamiliar with the formalism, and helpful for those who would like to use it as a starting point for a serious study of the field. To this end, we have tried to keep the presentation in the body of the text as intuitive as possible, while relegating technicalities and references to background literature to footnotes and an Appendix.[4]

[4] The main text presumes that the reader is familiar with the formalism of nonrelativistic

We begin in the next section with a sketch of the formalism of gauge theories, and a brief discussion of their interpretative problems. This is followed in §3 by a discussion of how general relativity itself may be cast as a gauge theory, and how in this context the hole argument can be viewed as a special case of the general interpretative problem of gauge invariance. In §4 we bring out some of the potential demerits of reading the general covariance of general relativity as a principle of gauge invariance. Most importantly, we discuss the fact that this reading seems to force us to accept that change is not a fundamental reality in classical and quantum gravity. This sets up the discussion of the following two sections, where we survey a number of proposals for understanding the general covariance of general relativity and discuss the proposals for quantizing gravity which they underwrite. Finally, in §7, we argue that the proposals canvassed in §§5 and 6 are directly related to interpretative views concerning the ontological status of the spacetime of general relativity. We conclude that problems about general covariance are indeed intimately connected with questions about the correct quantization of gravity, and the nature of time and change in physical theory.

2 Hamiltonian and Gauge Systems

There are a number of ways to formulate classical physical theories. One of the most straightforward is to proceed as follows. Construct a space whose points represent the physically possible states of the system in which you are interested. Then introduce some further structure which singles out a set of curves in this space which correspond to dynamically possible histories of the system. In the first two subsections we will sketch two implementations of this strategy: the Hamiltonian formalism and the gauge-theoretic formalism. We will see that the notion of a gauge system is a modest generalization of the notion of a Hamiltonian system—one simply weakens the geometric structure which is imposed on the space of states. As will become clear in the third subsection, however, this relatively small difference generates some very interesting interpretative problems: in the context of gauge systems, one is forced to make difficult decisions concerning the nature of the representation relation which holds between the mathematical space of states and the set of physically possible states of the system. We close the section with a brief discussion of the quantization of gauge systems.

2.1 Hamiltonian Systems

Many classical physical systems can be modeled by *Hamiltonian systems*. These are triples of mathematical objects, (M, ω, H). Here M is manifold, and ω is

quantum mechanics and with enough differential geometry to be able to read the standard textbook presentations of general relativity. The most important technical details are collected together in an Appendix. Although we hope that our presentation is not misleading, it does of course leave out many details.

a tensor, called a *symplectic form*, which gives M a geometric structure. The pair (M, ω) is called a *symplectic geometry*; the dension of M, if finite, must be even.. For our purposes, it is sufficient to note two ways in which the symplectic structure ω interacts with the set $C^{\infty}(M)$ of smooth real-valued functions on M. The first is that the symplectic structure ω gives us, *via Hamilton's equations*, a map $f \mapsto X_f$ between smooth functions on M and vector fields on M. Given $f \in C^{\infty}(M)$, one can integrate its vector field, X_f, to obtain a unique curve through each point of M (that is, one looks for the family of curves whose tangent vector at $x \in M$ is just $X_f(x)$). Thus we can associate a partition of M into curves with each smooth function on M. The second, and related, important function of the symplectic structure is to endow the set $C^{\infty}(M)$ with an interesting algebraic structure, the *Poisson bracket*. This is a binary operation which associates a smooth function, denoted $\{f, g\}$, with each pair of functions $f, g \in C^{\infty}(M)$. Intuitively, $\{f, g\}$ measures the rate of change of g along the set of curves generated by f, so that g is constant along the curves generated by f iff $\{f, g\}$=0. The Poisson bracket plays a crucial role in quantization.

We construct a Hamiltonian system by supplementing a symplectic geometry (M, ω) by a choice of a distinguished element $H \in C^{\infty}(M)$, called the *Hamiltonian*. The set of curves on M determined by ω and H are called the *dynamical trajectories*. Figure 1 depicts a Hamiltonian system: at the top, we have a symplectic geometry (M, ω); specifying a Hamiltonian serves to determine a unique dynamical trajectory through each point.

Taken together, H and (M, ω) constitute a theory of the behavior of the system in the following sense. We think of (M, ω) as being the space of dynamically possible states of some physical system—the *phase space* of the system. Each point of (M, ω) corresponds to exactly one physically possible state of the system, so a curve in phase space corresponds to a history of physically possible states of the system. To say that there is a unique dynamical trajectory through each point is to say that our theory specifies a unique past and future for every possible present state of the system. It is a complete and deterministic theory.

In the context of classical mechanics, one typically constructs a phase space by beginning with a smaller space, Q, the *configuration space* , which is taken to be the space of possible configurations of some set of particles or fields relative to physical space. One then identifies the phase space with the *cotangent bundle*, T^*Q, of Q. A point of T^*Q is a pair (q, p) where $q \in Q$, and p is a covector at q. If Q represents the set of possible positions of some set of particles relative to physical space, then T^*Q can be thought of as the space of possible positions and momenta of these particles. We can tell a similar story about fields. There is a canonical way of endowing T^*Q with a symplectic structure. We can now impose the Hamiltonian, H, whose value at a point (q, p) is just the energy of a system with that position and momentum. The dynamical trajectories for this Hamiltonian ought to model the observed behavior of our system.

Examples:

(1) The free particle. If we are dealing with a single particle in Euclidean space, then $Q = \Re^3$—the space of possible configurations of the particle is just

the space of positions of the particle relative to physical space. The phase space is $T^*Q = T^*\Re^3$, and H is just the kinetic energy. More generally, if we have a free particle moving in a physical space which is modeled by a Riemannian geometry, (S, g), then the configuration space is S and the phase space is T^*S endowed with the canonical symplectic form, ω. The dynamical trajectories corresponding to the particle moving along the geodesics of (S, g) are again generated by setting the Hamiltonian equal to the kinetic energy, $g^{ab}p_a p_b$.

(2) The Klein-Gordon Field. Fix a simultaneity slice in, Σ, Minkowski spacetime, and let Q be the space of configurations on this slice of the Klein-Gordon field ϕ of mass m—thus each point in Q corresponds to a $\phi : \Re^3 \to \Re$. We then look at T^*Q, where a point corresponds to a pair $(\phi, \dot{\phi})$. Our phase space consists of T^*Q equipped with the canonical symplectic structure, ω. Fixing an arbitrary timeslice, Σ, we can write:

$$\omega((\phi_1, \dot{\phi}_1), (\phi_2, \dot{\phi}_2)) = \int_\Sigma \phi_1 \dot{\phi}_1 - \phi_2 \dot{\phi}_2 \, dx^3$$

(here we are dealing with a linear field theory, so that Q and T^*Q are vector spaces, and we may identify vectors on Q with elements of T^*Q). For our Hamiltonian, we take:

$$H = \frac{1}{2} \int_\Sigma \left(\dot{\phi}^2 + \nabla\phi + m\phi^2 \right) \, dx^3$$

The equation of motion is just the usual Klein-Gordon equation,

$$\partial^a \partial_a \phi - m\phi = 0.$$

2.2 Gauge Systems

We consider an especially interesting generalization of the Hamiltonian framework: gauge systems.[5] The starting point is to relax one of the conditions imposed upon symplectic forms. This leads to a more general class of geometries, known as *presymplectic geometries*, which serve as the phase spaces of gauge theories. The presymplectic structure, σ, of a presymplectic geometry (N, σ) determines a natural foliation of the manifold N by submanifolds of some fixed dimension—there is one such submanifold through every point of N (see the top half of Figure 2). These submanifolds are called the *gauge orbits* of (N, σ).

Since (N, σ) is partitioned by gauge orbits, 'being in the same gauge orbit' is an equivalence relation. We denote this relation by $x \sim y$, and denote the gauge orbit of x by $[x]$. We call a diffeomorphism $\Phi : N \to N$ a *gauge transformation* if it preserves gauge orbits—i.e. if $x \sim x'$ implies $\Phi(x) \sim \Phi(x')$. We call a function $f : N \to \Re$ *gauge-invariant* if f is constant on each gauge orbit—i.e. if $x \sim x'$ implies $f(x) = f(x')$.

[5] The original development of the formalism, Dirac 1964, remains the best place to learn about gauge systems. Chapters 1–3 of Henneaux and Teitelboim 1992 include many invaluable examples as well as an introduction to the modern geometric point of view.

If we take a gauge-invariant function, H, on N as our Hamiltonian, then we can use the resulting *gauge system*, (N, σ, H), to model physical systems. We can again investigate the dynamical trajectories generated by H. Whereas in the Hamiltonian case there was a single dynamical trajectory through each point of phase space, we find in the gauge-theoretic case that there are infinitely many trajectories through each point.

The saving grace is that the family of dynamical trajectories through a given point, although they in general disagree radically about which point represents the future state of the system at a given time, do agree about which gauge orbit this point lies in. That is: if $x(t)$ and $x'(t)$ are dynamical trajectories which have their origin at the same point $x(0) = x'(0) = x_0$, then we have that $x(t) \sim x'(t)$ for all $t \in \Re$. Thus, although the presymplectic geometry is not strong enough to determine a unique dynamical trajectory through each point, it *is* strong enough to force all of the dynamical trajectories through a given point to agree about which gauge orbit the system occupies at a given time (cf. Figure 2). In particular, if f is a gauge-invariant function on N then the initial value problem for f is well-posed in the sense that if we fix an initial point $x_0 \in N$, then for any two dynamical trajectories $x(t)$ and $x'(t)$ which have their origin at the same point $x(0) = x'(0) = x_0$, we find that $f(x(t)) = f(x'(t))$ for all t. Thus specifying the initial state of the system completely determines the past and future values of any gauge-invariant quantity.

In practice, the most interesting gauge systems arise as *constrained Hamiltonian systems*. This means that our presymplectic phase space (N, σ) arises by restricting attention to a regular submanifold, N, of a symplectic geometry (M, ω), where N is equipped with the presymplectic form $\sigma = \omega\mid_N$ (the restriction of ω to N).[6] We introduce the notation $f \cong g$ (read "f is weakly equal to g") to indicate $f\mid_N = g\mid_N$, where $f, g \in C^\infty(M)$.

Locally, we can specify N by requiring that some set, $\mathcal{C} = \{C_a\}$ of real-valued functions on M vanish. Such functions are called *constraints*. There are two kinds of constraints: if $C_a \in \mathcal{C}$ is such that for $x \in N$, $X_{C_a}(x)$ is tangent to $[x]$, 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. Equivalently, the first class constraints are those which commute with all of the constraints. The first class constraints, but not the second class constraints, generate gauge transformations on N.[7] That is: following in N the integral curve of a vector field associated with a first class constraint carries one along gauge orbits of (N, σ) (here we are thinking of vector fields as the infinitesimal generators of diffeomorphisms). In fact, at each point $x \in N$, $\{X_{\gamma_a}(x)\}$ is a basis for the tangent space of $[x]$ so that the dimensionality of the gauge orbit of x is just the cardinality of $\{\gamma_a\}$.

A function $f : M \to \Re$ has a gauge-invariant restriction to N iff $\{f, \gamma\} \cong 0$ for all first class constraints γ (here $\{,\}$ are the Poisson brackets on (M, ω)). One describes this result by saying that gauge-invariant functions *commute* with

[6] A regular submanifold is one which is given locally by stipulating that some subset of a set of coordinates on M take on a given constant value.

[7] Here we gloss over the subtleties surrounding the Dirac conjecture. See pp. 16–20 of Henneaux and Teitelboim 1992.

the first class constraints. In what follows we shall be exclusively concerned with first class constraints.

Examples:

(3) Let (T^*Q, ω) be a finite dimensional cotangent bundle with its canonical symplectic structure, and let (q^i, p_i) be canonical coordinates. Now let (N, σ) arise by imposing the first class constraint $p_1 = 0$.[8] The Hamiltonian vector field of p_1 in (M, ω) generates motions in the q^1 direction, so the gauge orbits in (N, σ) are of the form $\{(s, q^2, \ldots; p_2, \ldots)$: where $s \in \Re$ and all other q^i and p_i are fixed$\}$. Thus the gauge-invariant functions on N are those which are independent of q^1. For any Hamiltonian, a dynamical trajectory is of the form $(q^i(t), p_i(t))$, with $q^i(t)$ and $p_i(t)$ determined uniquely for $i \geq 2$, but with $q^1(t)$ an arbitrary function of time—we call the physically irrelevant q^1 a *gauge* degree of freedom. 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 local canonical coordinates, $(q^i; p_i)$, on (M, ω) so that the first class constraints are of the form $p_i = 0$ for $i \leq k$ and the $q^i(t)$ are arbitrary for $i \leq k$.

(4) Vacuum Electromagnetism. Let (S, g) be a three dimensional Riemannian manifold representing physical space. And let $Q = \{A : S \to \Re^3\}$ be the infinite dimensional space of covector fields on S—that is, each element of Q is a function which maps each point of S to a three-vector. We construct the cotangent bundle, T^*Q. A point in T^*Q is a pair (A, E), where E, like A, is a vector field on S (again, identifying vectors on a linear space with elements of that space). We endow T^*Q with the canonical symplectic structure, ω. In order to construct the phase space of electromagnetism, we restrict attention to those points $(A, E) \in T^*Q$ such that div $E = 0$. This is a first class constraint. The constraint manifold, N, is an infinite dimensional submanifold of T^*Q. We equip N with the presymplectic form, $\sigma = \omega \mid_N$. The presymplectic manifold, (N, σ) is the phase space of electromagnetism. The gauge orbits of (N, σ) are determined by the following equivalence relation: $(A, E) \sim (A', E')$ iff $E' = E$ and $A' = A + \text{grad}\ \Lambda$ for some $\Lambda : S \to \Re$. Thus, $[(A, E)] = \{(A', E) : A' = A + \text{grad}\ \Lambda,\ \Lambda : S \to \Re\}$. What are the gauge-invariant functions on this phase space? If, e.g., we fix a point $\xi \in S$, then the function $\xi_E : N \to \Re^3$ whose value at (A, E) is just $E(\xi)$ is gauge-invariant. On the other hand, the function which returns the value $A(\xi)$ is clearly *not* gauge-invariant (in general, $A(\xi) \neq A'(\xi)$ even if $(A, E) \sim (A', E)$). We can, however, use A to construct gauge invariant quantities. Of these, the most important is the magnetic field, $B \equiv \text{curl}\ A$. Since curl $(A) =$curl $(A+\text{grad}\ \Lambda)$ for any scalar Λ, we find that $B(\xi) = B'(\xi)$ whenever $(A, E) \sim (A', E)$.

We choose our Hamiltonian to be $H = \int_S (\mid E \mid^2 + |\text{curl}\ A\mid^2)\, dx$. Hamilton's equations are $\dot{A} = -E$ and $\dot{E} =$curl (curl A). 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

[8] Note that if we imposed both $p_1 = 0$ and $q^1 = 0$, then the constraints would be second class, and (N, σ) would be a symplectic geometry.

not serve to determine a unique dynamical trajectory. But we do find that if $(A(t), E(t))$ and $(A'(t), E'(t))$ are solutions of Maxwell's equations for the initial data $(A_0, E_0) \in N$, then for each t, $E(t) = E'(t)$ and there is a scalar function on space, $\Lambda(t)$, such that $A'(t) = A(t) + \text{grad } \Lambda(t)$. Equivalently: if $(A(t), E(t))$ and $(A'(t), E'(t))$ are dynamical trajectories with their origins in the same point of (N, σ), then we have that $[(A(t), E(t)] = [(A'(t), E'(t)]$ for all t. Maxwell's equations do not determine the future value of $A(t)$, but they do determine in which gauge orbit $A(t)$ will lie.

2.3 Interpreting Gauge Theories

The interpretation of theories cast in Hamiltonian form is typically quite straightforward. Given a Hamiltonian system, (M, ω, H), one can always stipulate that it represents a system whose dynamically possible states stand in a one-to-one correspondence with the points of M (call this the *literal* approach to interpreting the theory). Furthermore, in the context of classical mechanics, it often happens that $M = T^*Q$, where Q can be viewed as the space of possible configurations of a set of particles or fields relative to some inertial frame. In this case, one ends up with an extremely attractive interpretation of the theory as a deterministic account of a physically reasonable system.

Unfortunately, interpreting gauge theories is seldom so simple. In the case of electromagnetism, the application of the literal strategy leads to the claim that our system has a distinct dynamically possible state for every pair (A, E). But then one is committed to viewing electromagnetism as an indeterministic theory: specifying the initial dynamical state, (A_0, E_0), fails to determine the future dynamical state, since if (A, E) is a dynamically possible state at time t according to Maxwell's equations, then so is $(A + \text{grad } \Lambda, E)$. The present state of the electromagnetic field fails to determine the future state of the field. Clearly, the same sort of indeterminism will arise whenever a gauge theory is given a literal interpretation.

This flies in the face of common sense: given initial data one can use Maxwell's equations to make highly accurate predictions. So there has to be something wrong with our literal interpretation of the theory. There are two possible diagnoses here. The first is that the interpretation, although essentially correct, needs to be supplemented with an account of measurement which will insure that the predictions derivable from our gauge theory are perfectly determinate. The second is that the formalism of our gauge theory presented above contains "surplus structure," which must be eliminated—either at the level of formalism or the level of interpretation—if we are to have a physically sensible understanding of the theory.[9] We consider each of these alternatives in turn.

If we wish to stick with our literal interpretation of our gauge theory, then we have to explain how it is that the theory is used to make determinate predictions despite its indeterminism. The most obvious way of doing so is to claim

[9]See Redhead 1975 for the notion of surplus structure, and its relevance to the interpretative enterprise.

that some physically real quantities are not measurable. In order to produce determinate predictions, we need to work with physical quantities whose initial value problems are well-posed. A function on phase space has a well posed initial value problem iff it is gauge-invariant. So we will want to stipulate that only gauge-invariant quantities are measurable. This allows us to maintain predictability, even in the face of indeterminism. In the case of electromagnetism, implementing this strategy will mean accepting gauge-invariant quantities like the electric field, E, and the magnetic field, B, as measurable, while denying the vector potential, A, is directly measurable. Nonetheless, the vector potential will be a physically real quantity: since every point of phase space corresponds to a distinct physically possible situation, (A, E) and (A', E') will represent distinct situations in virtue of disagreeing as to the value of the quantity A—even if $[A] = [A']$ so that the two states of affairs are not empirically distinguishable.

This sort of ploy is likely to seem rather desperate, however. It seems far more natural to insist that the *only* physically real quantities are gauge-invariant quantities (call this strategy the adoption of a *gauge-invariant* interpretation). In this case, one needn't resort to a tricky account of measurement: one can stick to the orthodox position that every classical physical quantity is (in principle) measurable with arbitrary accuracy. Furthermore, the interpretation renders the theory deterministic, since specifying the initial state determines the future and past values of the physically real quantities. In the case of electromagnetism, for instance, it is natural to maintain that E and B taken together encode all of the structure of the electromagnetic field. When physical space is simply connected, the divergence free magnetic fields are in one-to-one correspondence with the gauge orbits of vector potentials, so that this move is tantamount to taking $[A]$ rather than A as the physically real quantity. The resulting interpretation is fully deterministic, and supports an orthodox account of measurement. Notice that this establishes that determinism cannot be a formal property of theories: to ask whether electromagnetism is deterministic or not is not to ask a technical question about the formalism of example (4); rather it is to ask whether one prefers a literal or gauge-invariant interpretation of this formalism.

There is a formal move which is associated with the interpretative move from literal to gauge-invariant interpretations: *reduction*. As it stands, our formalism is good at predicting which gauge orbit we will end up in, but lousy at predicting which point we will end up at. This suggests that what we really need is a theory of gauge orbits rather than points. Thus we attempt to do the following: we build a new manifold, $\tilde{M}$, whose points are the gauge orbits of (N, σ); we then use σ to construct a form $\tilde{\omega}$ on $\tilde{M}$; finally, we use H to induce a Hamiltonian $\tilde{H}$ on $(\tilde{M}, \tilde{\omega})$. This is called the *reduced phase space*. It is not always possible to carry out this construction: the set of gauge orbits will not be a manifold, even locally, unless there exist sufficiently many gauge-invariant quantities to fix a gauge orbit. If not, it will of course be impossible to construct a symplectic form on $\tilde{M}$. But when $\tilde{M}$ is well behaved, $\tilde{\omega}$ is a symplectic form so that $(\tilde{M}, \tilde{\omega}, \tilde{H})$ is a genuine Hamiltonian system rather than a gauge system. This Hamiltonian system describes the way in which the trajectories of the original gauge system travel through gauge orbits. Giving a gauge-invariant interpretation of the original gauge theory is the same thing as giving a literal interpretation of the reduced phase space. In a sense, then, it is always easy to find a gauge-invariant interpretation, barring technical difficulties: simply construct the reduced phase space and adopt a literal interpretation. It can happen, however, that the reduced phase space doesn't admit any physically attractive literal interpretations—it needn't, for instance, have the structure of a cotangent bundle over configuration space for reasonable particles or fields. Something like this actually happens in the case of electromagnetism when space is multiply connected.[10]

Gauge systems differ from Hamiltonian systems in that their equations of motion fail to determine the evolution of all of their variables. In the classical context, it is reasonable to regard this fact as reflecting a shortcoming of the formalism (the inclusion of excess variables) rather than a genuine ontological indeterminism. That is: it is preferable to look for interpretations in which only those variables whose evolution is determined by the equations of motion are taken to correspond to physically real quantities. If we can find a large enough set of such quantities to fix the gauge orbit of the system, and which can be taken to correspond to plausible physical quantities, then we have found an acceptable interpretation. In the case of ordinary vacuum electromagnetism on a simply connected spacetime, E and B serve this function admirably. In what follows, we will see that many foundational issues in classical and quantum gravity turn upon the difficulty of finding a complete set of physically reasonable gauge-invariant quantities for general relativity.

2.4 Quantizing Gauge Theories

A quantization of a Hamiltonian system (M, ω, H) consists of a Hilbert space, $\mathcal{H}$, equipped with a Hamiltonian operator, $\hat{H}$, and a representation of an appropriate subalgebra of the Poisson algebra of classical observables as an algebra of self-adjoint operators on $\mathcal{H}$. If M can be written as T^*Q for some natural configuration space Q , then one normally chooses $\mathcal{H}$ to be $L^2(Q, \mu)$, the space of complex functions on Q which are square-integrable with respect to some physically relevant measure μ.

How does one quantize a gauge theory? There are two main routes. The first is to construct the reduced phase space and apply canonical technique to the resulting Hamiltonian system. This tends to be impracticable, however—even when the reduced phase space exists, its structure is often difficult to determine. The alternative is to quantize the gauge system directly, employing a technique due to Dirac.[11]

[10] When space is multiply connected, the correspondence between gauge orbits and magnetic fields in many-to-one. This means that there is additional structure that is not captured by B—structure which is empirically accessible via the Aharonov-Bohm effect. See Belot 1998 for an account of how this complication forces the would-be interpreter of electromagnetism to choose between non-locality and indeterminism. The moral is that it is not always easy to find a gauge-invariant interpretation of a given gauge theory.

[11] Again, Dirac 1964 remains one of the best sources. See also Chapter 13 of Henneaux and Teitelboim 1992. The following description, although accurate enough for present pur-

Suppose that one has a gauge system (N, σ, H) where (N, σ) is the submanifold of a symplectic geometry (M, ω) determined by the first class constraint γ given by $C \equiv 0$. Then one chooses a set of coordinates on M, and finds a vector space, V, which carries a representation of their Poisson algebra as linear operators: if (p, q) are canonical coordinates on M, then one looks for operators $\hat{q}$ and $\hat{p}$ satisfying $[\hat{q}, \hat{p}] = -i\hbar$. One then looks for a quantum analog, $\hat{C}$, of the classical constraint—e.g., if $C = p^2$ then $\hat{C} = \hat{p}^2$.[12] Next, one imposes the quantum constraint to construct the space of physical states: $V_{phys} = \{\psi \in V : \hat{C}\psi = 0\}$. This ensures that the quantum states are gauge-invariant: if a given degree of freedom, q, is gauge (i.e. physically irrelevant) at the classical level, then it should be gauge at the quantum level.

Suppose, for example, that the classical constraint is $C = p$. Then we know from example (3) that the classical degree of freedom, q, is gauge. Thus gauge-invariant functions are independent of q. Working in the standard Schrödinger representation, we have that

$$\hat{C} = \hat{p} = i\frac{\partial}{\partial q}$$

so that stipulating that $\hat{C}\psi = 0$ amounts to requiring that the quantum wave functions be independent of the gauge degree of freedom, q. Similarly, imposing the quantum constraint corresponding to div $E = 0$ forces the states $\psi(A)$ of quantum electrodynamics to be independent of the choice of gauge (i.e. if $A' = A + \text{grad}\, \Lambda$, then $\psi(A) = \psi(A')$).

Finally, one looks for an appropriate inner product to make V_{phys} into a Hilbert space, and for an appropriate quantum Hamiltonian, $\hat{H}$, which determines the quantum dynamics via Schrödinger's equation.

3 General Relativity as a Gauge Theory

In its standard version, the hole argument looks something like the following (see Earman and Norton 1987). Let $\mathcal{M} = (M, g)$ be a model of general relativity, and let $d : M \to M$ be a diffeomorphism (called a *hole* diffeomorphism) which differs from the identity only on some small open set, U. The general covariance of the theory implies that $\mathcal{M}' = (M, d^*g)$ is also a model. If one views $\mathcal{M}$ and $\mathcal{M}'$ as representing distinct physically possible worlds, then one is committed to believing that general relativity is an indeterministic theory—specifying the state of the gravitational field on a Cauchy surface prior to U fails

poses, glosses over a large number of technicalities. One would hope, of course, that the two techniques of quantization would lead to the same results in cases where they can both be carried out. Unfortunately, this is not always the case; see Plyuschay and Razumov 1995. See Landsman 1995 and 1998 for an alternative to Dirac quantization.

[12] Note that operator ordering problems introduce considerable ambiguity at this stage. We ignore these below. In general, it is safe to assume that quantum gravity is beset by all of the problems of ordinary quantum field theories—operator ordering problems, divergences, anomalies, problems of renormalization and regularization—and then some.

to determine the state of the field inside U. Furthermore, it is claimed, if one is a substantivalist about the spacetime points of general relativity, then, *prima facie*, one *is* committed to viewing $\mathcal{M}$ and $\mathcal{M}'$ as representing distinct states of affairs. The conclusion is that substantivalists are *prima facie* committed to the doctrine that general relativity is an indeterministic theory.

In this section we will show that the hole argument is a special case of the observation made in the previous section: a gauge theory is indeterministic under a literal interpretation. We sketch the formulation of general relativity as a gauge theory and then argue that certain forms of substantivalism are, in fact, literal interpretations of this formalism.

3.1 Formalism

We begin our search for a gauge-theoretic formulation of general relativity 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 Einstein field equations, (M, g), whose Cauchy surfaces are diffeomorphic to Σ.[13] We embed Σ in (M, g) via a diffeomorphism $\phi : \Sigma \to M$ such that $S = \phi(\Sigma)$ is a Cauchy surface of (M, g), and we study the geometry which g induces on S. We will take this geometry to represent an instantaneous state of the gravitational field.[14] This geometry is characterized by two symmetric tensors on S, q_{ab} and K_{ab}. Here 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

[13]We limit discussion to globally hyperbolic vacuum solutions of general relativity 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 largely for convenience's sake. Much 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. But, for the most part, taking these differences into account would involve adding many qualifications to our technical treatment, without substantially altering the interpretative theses defended below.

[14]There are two ways to proceed here. We follow that more familiar route, and characterize the geometry of S using a metric tensor (geometrodynamics). It is also possible to work in terms of connections (connection dynamics). This approach, pioneered by Ashtekar, is in many ways more tractable and has led to many significant results in recent years. The best intuitive introductions to connection dynamics are contained in Baez and Munian 1994 and Kuchař 1993a; see Ashtekar 1995 for a more detailed presentation. We believe that at the level of detail of the present paper, nothing is lost by focusing on geometrodynamics to the exclusion of connection dynamics.

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. 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 some (M, g) in such a way that q and K arise as the first and second fundamental forms of Σ iff the following two relations—known as the *Gauss* and *Codazzi* constraints respectively—hold:

$$R + (K^a{}_a)^2 - K^{ab}K_{ab} = 0$$

$$\nabla^a K_{ab} - \nabla_b K^a{}_a = 0.$$

Here the metric q on Σ is used to define the scalar curvature, R, and the covariant derivative, ∇. Note that these conditions make reference only to q and K—they do not mention g.

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. The natural starting point for writing down general relativity 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 then endow it with the canonical symplectic structure, ω. The momentum canonically conjugate to q is given not by K but by

$$p^{ab} \equiv \sqrt{\det q}(K^{ab} - K^c{}_c q^{ab}).$$

The phase space of general relativity is the constraint surface $N \subset T^*Q$ given by the following first class constraints, known as the *scalar* and *vector* constraints (or, alternatively, as the *Hamiltonian* and *momentum* constraints):

$$h \equiv \sqrt{\det q}(p^{ab}p_{ab} - \frac{1}{2}(p^a_a)^2 - R) = 0$$

$$h_a \equiv \nabla_b p^b{}_a = 0.$$

Each of these equations actually 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 the Gauss and Codazzi constraints, rewritten in terms of p rather than K. Let $\sigma = \omega \mid_N$, and let $H \equiv 0$. Then general relativity is the gauge theory (N, σ, H).[15]

[15] See Appendix E of Wald 1984 for a more complete treatment. Note that most formulations of general relativity as a gauge theory make use of the lapse and shift as Lagrange multipliers. In order to avoid this complication, we have followed Beig 1994 in adopting a more geometric approach in which the lapse and shift are eliminated.

At each point $x \in N$, the gauge orbit of (N, σ, H) is infinite dimensional. These orbits have the following structure. Fix $x = (q, p)$ and $x' = (q', p')$ in N. Then x and x' lie in the same gauge orbit iff there is a solution to the Einstein field equations, (M, g), and embeddings, $\phi, \phi' : \Sigma \to M$, 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 describe spatial geometries of the *same* model of general relativity. Thus each gauge orbit can be viewed as being the space embeddings 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 embeddings for any other model isometric to (M, g)). This means that each dynamical trajectory lies in a single gauge orbit: as the gravitational field evolves, it always stays in the same gauge orbit. This is, in fact, the significance of setting $H \equiv 0$: the vanishing of the Hamiltonian means that the dynamical trajectories are always tangent to the gauge orbits, which is just to say that once a dynamical trajectory is in a given gauge orbit, it never leaves. As we will see below, a zero Hamiltonian is closely related to the lack of a preferred time parameter.

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 M by Cauchy surfaces (which are, of course, all diffeomorphic to Σ). We then choose a time function $\tau : M \to \Re$, 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 $\Phi : M \to \Sigma \times \Re$ such that each Cauchy surface of the foliation, S, is mapped onto a set of the form $\Sigma \times \{t\}$. 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 push forward g, so that $\mathcal{M}' = (\Sigma \times \Re, \Phi^* g)$ 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 $p^{ab}(t)$ characterize the geometry of the Cauchy surface $\Sigma \times \{t\}$ in $\mathcal{M}'$. As t varies, $(q_{ab}(t), p^{ab}(t))$ sweeps out a curve in N —the points on this curve representing a sequence of Cauchy surfaces. This curve is a dynamical trajectory of (N, σ, H). Choosing a different foliation, time function or identification map gives us a new dynamical trajectory, which will be related to the first by a gauge transformation—i.e., one can map one dynamical trajectory on to the other via a transformation of phase space which preserves gauge orbits.

The trajectories which correspond to the models $\mathcal{M}$ and $\mathcal{M}'$ which appear in the hole argument are so related. This shows that our approach respects the general covariance of general relativity in the sense that it is indifferent to changes of foliation, time function, and identification map. Changing any of these simply carries us from one dynamical trajectory to a gauge related one.[16]

[16] Actually, this glosses over an interesting detail: it could be argued that our formalism *fails* to be diffeomorphism invariant, since our phase space only contains *spacelike* geometries. This is closely related to the fact, emphasized to us by Steve Weinstein, that it is far from trivial to see how the group of four dimensional diffeomorphisms acts on the phase space of three-geometries. See Kuchař 1986, Isham 1991, and Weinstein 1998 for illuminating discussions of

Now suppose that we look at two points $x = (q, p)$ and $x' = (q', p')$ 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 $d : \Sigma \to \Sigma$ such that $d^*q = q'$ and $d^*p = p'$. 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. *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 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 is generated by a combination of both sorts of constraint.

Next, 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).

Modulo technical difficulties to be discussed in §4, we can convert this gauge theory into a true Hamiltonian system by factoring out the action of the gauge transformations to construct the reduced phase space. It is illuminating to proceed in two steps: we first partially reduce the phase space by factoring by the action of the gauge transformations generated by the vector constraint; we then complete the reduction by removing the gauge freedom associated with the scalar constraint.

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 $Q =$Riem(Σ) of metrics on Σ by $Q_0 =$Riem(Σ)/Diff(Σ), the set of equivalence

this problem in the classical and quantum theories. The results announced in Gotay, Isenberg, and Marsden 1998 promise to shed a great deal of light on this problem.

classes of diffeomorphically related metrics on Σ. We call Q_0 *superspace*. We now construct T^*Q_0 and impose the scalar constraint, to construct the presymplectic geometry $(\tilde{N}, \tilde{\sigma})$. The gauge theory $(\tilde{N}, \tilde{\sigma}, H \equiv 0)$ is the partially reduced phase space formulation of general relativity. 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. If we now identify points in $\tilde{N}$ which are related by a gauge transformation generated by the scalar constraint, then we end up with a Hamiltonian system $(\tilde{M}, \tilde{\omega}, H \equiv 0)$, where points in the phase space correspond to equivalence classes of diffeomorphically related models of general relativity.

3.2 Interpretation

Classical substantivalists and relationalists about space agree with one another that space exists and that it has some fixed geometrical structure. They are divided over the question of the nature of the existence of this peculiar entity. Substantivalists hold that it consists of parts which maintain their identity over time, and that these parts stand directly in geometrical relations to one another, while material objects stand in spatial relations only in virtue of the relations obtaining between those parts of space which they occupy. Relationalists deny the substantivalist claim that space has genidentical parts. They maintain that space is best thought of as the structure of possible spatial relations between bodies. Such relations are to be taken as primitive, rather than being reduced to relations holding among the points of an underlying substratum. There are a couple of vivid ways of putting the issue between the two factions. Substantivalists, but not relationalists, believe in genidentical points. This means that substantivalists, but not relationalists, can help themselves to a straightforward account of the nature of absolute motion—it is motion *relative* to the genidentical parts of space.[17] In addition, substantivalists will follow Clarke in affirming, while relationalists will follow Leibniz in denying, that two possible worlds could instantiate all of the same spatial relations, but differ in virtue of which point of space plays which role (I occupy *this* point rather than *that* one).

Relativistic physics, however, seems to demand that one think in terms of spacetime rather than space. Thus, the traditional doctrines are often translated into the four dimensional context. Substantivalists and relationalists will again agree that the world has some given geometrical structure. Substantivalists understand the existence of spacetime in terms of the existence of its pointlike parts, and gloss spatiotemporal relations between material events in terms of the spatiotemporal relations between points at which the events occur. Relationalists will deny that spacetime points enjoy this robust sort of existence, and will accept spatiotemporal relations between events as primitive. It is now somewhat more difficult to specify the nature of the disagreement between the

[17] There are, however, some more sophisticated ploys which relationalists can adopt to make sense of inertial effects. See Barbour 1982, Belot 1999, and Lynden-Bell 1995.

two parties. It is no longer possible to cash out the disagreement in terms of the nature of absolute motion (absolute acceleration will be defined in terms of the four dimensional geometrical structure that substantivalists and relationalists *agree* about). We can, however, still look to *possibilia* for a way of putting the issue. Some substantivalists, at least, will affirm, while all relationalists will deny, that there are distinct possible worlds in which the same geometries are instantiated, but which are nonetheless distinct in virtue of the fact that different roles are played by different spacetime points (in this world, the maximum curvature occurs at *this* point, while it occurs at *that* point in the other world). We will call substantivalists who go along with these sort of counterfactuals *straightforward* substantivalists. Not all substantivalists are straightforward: recent years have seen a proliferation of *sophisticated* substantivalists who ape relationalists' denial of the relevant counterfactuals (see Brighouse 1994, Butterfield 1989, Field 1985, and Maudlin 1990). For the time being, however, we will bracket this option. We will address the virtues and vices of sophisticated substantivalism in §7.

It is easy to see that (straightforward) substantivalists are committed to giving a literal interpretation of general relativity. Consider two models, $\mathcal{M} = (M, g)$ and $\mathcal{M}' = (M, d^*g)$, which are related by a hole diffeomorphism, d. Fix a foliation, time function, and identification map, and use them to construct dynamical trajectories $x(t)$ and $x'(t)$ in the phase space of general relativity which correspond to $\mathcal{M}$ and $\mathcal{M}'$. Because d is a hole diffeomorphism, we can assume that $x(t) = x'(t)$ for $t \leq 0$, but $x(1) \neq x'(1)$. Substantivalists will view $x(t)$ and $x'(t)$ as representing distinct physically possible histories: although they represent the same spatiotemporal geometry (lying as they do in the same gauge orbit), they represent different distributions of their shared set of geometrical properties over the points of spacetime (if x is a point on the spacelike surface $t = 1$, then x represents it as having *these* properties while x' represents it as having *those*). Indeed whenever x and x' are distinct points in the phase space of general relativity, a substantivalist will view them as representing distinct physical situations: either they represent distinct possible geometries for a given spacelike hypersurface, or they represent the same pattern of geometric relations differently instantiated. This is just to say that substantivalists are committed to a literal construal of the gauge-theoretic formulation of general relativity. And, like any literal interpretation of a gauge theory, substantivalism implies that the theory is indeterministic: if $x(1)$ and $x'(1)$ correspond to distinct possible situations, then the state corresponding to $x_0 = x(0) = x'(0)$ has multiple physically possible futures. This is the content of the hole argument.

As in §2, the best way to avoid this sort of indeterminism is to adopt a gauge-invariant interpretation of the theory. We can do this by giving a literal interpretation of the reduced phase space formulation of general relativity. Recall from above that the points of the reduced phase space are just the equivalence classes of diffeomorphic models of general relativity. Thus in order to avoid the indeterminism of the hole argument, we have to accept that diffeomorphic models always represent the same physically possible situation (this proposition is known as *Leibniz equivalence* in the literature on the hole argument). And this, of course, is just to deny that there could be two possible worlds with the same geometry which differ only in virtue of the way that this geometry is shared out over existent spacetime points. Thus, modulo the existence of an attractive form of sophisticated substantivalism, one must be a relationalist in order to give a deterministic interpretation of general relativity.

There is another, closely related, motive for adopting a gauge invariant interpretation of general relativity. As was noted at the end of §2, the existence of gauge degrees of freedom in a theory seems to tell us that the theory contains excess variables. The natural response is to seek an interpretation in which all and only the variables which correspond to physical degrees of freedom are taken seriously. Typically, we will want to say that it is just those variables whose evolution is determined by the differential equations of the theory that should be taken seriously in this way. Recently, a number of philosophers have joined the majority of physicists in advocating such gauge-invariant interpretations of general relativity—although almost all philosophers opt for a form of sophisticated substantivalism, while many physicists adhere to a strict relationalism.

At this point a potential technical problem looms. Relatively little is presently known about the structure of the reduced phase space of general relativity. It is known that this space has singularities corresponding to models of general relativity with symmetries, and is smooth elsewhere (Marsden 1981). Interesting and extensive smooth open sets have been constructed (Fischer and Moncrief 1996; see fn. 25 below). But the concern is sometimes expressed that the structure of generic regions of this space may not be smooth (see p. 141 of Kuchař 1993a, p. 267 of Unruh 1991, and p. 2600 of Unruh and Wald 1989). Equivalently, one can wonder whether there exists a full set of gauge-invariant quantities on the unreduced phase space of general relativity. In fact, very few such quantities are known (see Goldberg et al 1992 for a rare example). Furthermore, it is known that there are *no* local gauge-invariant quantities.[18]

Until some progress is made on these technical questions, a dark cloud hangs over the program of providing gauge-invariant interpretations of general relativity. The problem is this. One knows that the reduced phase space of general relativity exists as a mathematical set with some topology (although this topology may not be well enough behaved to support any interesting global geometric structure). And one knows that one can characterize the points of the reduced phase space as equivalence classes of models of general relativity. Philosophers who have advocated gauge-invariant interpretations have been satisfied with this sort of approach, which we dub *extrinsic*, since the characterization of the points of the reduced phase space is in terms of the gauge orbits of the original phase space. Such an extrinsic approach may, indeed, yield some sort of interpretation of general relativity. But we feel that something is lacking from an interpretation which stops at this point. Ideally, one would like an interpretation

[18] See Torre 1993. Here a quantity is local if it is an integral over Σ of the canonical variables, p and q, and a finite number of their derivatives. The situation is slightly more encouraging if we work with asymptotically flat spacetime, rather than spatially compact spacetimes. In that case, their are a finite number of known local gauge-invariant quantities, such as the ADM momenta.

of general relativity which was underwritten by some *intrinsic* characterization of the points of reduced phase space. Indeed, in order to formulate a gauge invariant quantum theory, one would like to be able to find a set of coordinates on the reduced phase space—or, equivalently, a full set of gauge invariant quantities on the original phase space. This would amount to isolating the true (i.e., gauge invariant) degrees of freedom of the theory. Although this is *not* essential for Dirac quantization, it nonetheless seems to us that it is the approach to the theory which yields the deepest understanding, since it underwrites an explicit characterization of the classical and quantum degrees of freedom of the system.[19]

Thus we conclude that the present state of ignorance concerning the structure of the reduced phase space of general relativity—and the lingering worry that this structure may be monstrous—should give pause to advocates of gauge-invariant interpretations of the theory. We will, however, bracket this technical objection to gauge invariant interpretations, and move on to discuss the two other sorts of problem which plague such interpretations.

(1) It appears to be a consequence of any gauge-invariant interpretation of general relativity that change does not exist, since any such interpretation requires us to regard two points, x and x', of the phase space of general relativity which correspond to distinct Cauchy surfaces of the same model as representing the same state of affairs, since they are related by a gauge transformation generated by the scalar constraint. Equivalently, if the only physical quantities are gauge-invariant, then there is no such quantity which allows us to distinguish between two such Cauchy surfaces.

(2) Accepting a gauge-invariant interpretation of general relativity, and thus treating the general covariance of general relativity as analogous to the gauge invariance of electromagnetism, leads to nasty technical and interpretative problems when one attempts to quantize the theory. These problems are so intractable that some have called for a re-evaluation of the standard understanding of general covariance.

We will discuss these problems for advocates of gauge-invariant interpretations in §§4 and 5. In §6, we will survey some interpretations which lie outside of the gauge-invariant orthodoxy. All of these options will be seen to have serious shortcomings, as well as distinctive attractive features.

4 Gauge Invariance and Change

Is there room for time or change when general covariance is understood as a principle of gauge invariance? *Prima facie*, a gauge-invariant interpretation of general relativity is descriptively inadequate because it cannot accommodate

[19] This is, we believe, part of the explanation of the current vogue for non-local interpretations of electromagnetism and its nonabelian cousins; see Baez and Munian 1994 for a nice introduction to these issues. It also justifies the demand, expressed in Earman 1989, that relationalists should produce formulations of physical theories which can be expressed in relationally pure vocabulary.

real change.

> To maintain that the only observable quantities are those that commute with all the constraints [i.e. the gauge invariant quantities] seems to imply that the Universe cannot change. For this reason, this standpoint on observables was dubbed the *frozen time formalism*. The frozen time formalism never successfully explained the evolution we see all around us. (Kuchař 1992, p. 293)
>
> How can changes in time be described in terms of objects which are completely time independent? In particular, since the only physical, and thus measurable quantities are those which are time independent, how can we describe the rich set of time dependent observations we make of the world around us? ... The time independent quantities in General Relativity alone are simply insufficient to describe time dependent relations we wish to describe with the theory. (Unruh 1991, p. 266)

Kuchař and Unruh are putting their fingers on an important question about the nature of time. It will be helpful in what follows to be clear on the relation between their question and questions about the nature of time which are currently at the center of philosophical discussion. Kuchař and Unruh are *not* interested in: (i) the direction of time; (ii) the objectivity of the metric structure of time; (iii) the reducibility of temporal relations to causal relations; or (iv) the existence of a moving now or flow of time. Rather they are interested in whether or not change itself exists. And, of course, to the extent that the existence of time and that of change are closely related, they are interested in the existence, or lack thereof, of time as well. Thus, it is tempting to see them as engaging the same problematic about time, change, and flux that so occupied the Ancients.

In the quotations above, Kuchař and Unruh are driving at the following point. If we accept that the only physically real quantities of general relativity are gauge-invariant, then it follows that for any given model there is no physically real quantity which takes on different values when evaluated on Cauchy surfaces corresponding to distinct times. Which is, they claim, just to say that there is no change when general relativity is understood in this fashion: there is no evolution in time of the values of the physically real quantities. *Prima facie*, people who hold such a view have a very simple view of the nature of change: it is illusory. For this reason, Kuchař associates the reading of the general covariance of general relativity as a principle of gauge invariance with the name of Parmenides (1993a, p. 139).

Both Kuchař and Unruh denounce this Parmenidean view. They maintain that it flies in the face of our experience of time and change, and are skeptical that any coherent conceptual framework for the articulation of a quantum theory of gravity can be built upon such a foundation. Many physicists working on quantum gravity seem to be swayed by these arguments. But if the Parmenidean view of change is to be rejected as descriptively inadequate, what sort of account should be erected in its place? This is clearly a philosophical problem. Now, the vast majority of contemporary philosophical discussion about the nature of change is concerned with the existence of a moving now. This literature, whatever its merits as metaphysics, seems to be entirely irrelevant to the physical problems with which we are here concerned—since it almost always presupposes a pre-relativistic world view, and turns upon a question (the viability of the tenseless view of time) which is likely to appear long-since settled to relativistically-minded physicists.

Philosophers come closest to the physicists' questions when they attempt to motivate the idea of a moving now. In introducing philosophical theses related to this latter problem, Le Poidevin and MacBeath comment that:

> It is a commonplace that time, not space, is the dimension of change. There is a wholly uncontroversial sense in which this is true: genuine change involves temporal variation in the ordinary properties of things: a hot liquid cools, a tree blossoms, an iron gate rusts. Purely *spatial* variation, for example the distribution of colours in patterned rug, does not count as a genuine change. Uncontroversial as this is, it requires explanation. What is special about time? (1993, p. 1)

In fact, it is not uncommon to introduce the moving now as a solution to this problem, before going on to consider whether it is a coherent notion (see, e.g., Mellor 1993, p. 163). Unfortunately, philosophers seem to have all too little to say about what distinguishes change from mere variation.

Unruh, however, has made a very interesting and influential suggestion along these lines—one which is clearly motivated by physical concerns, but which strikes us as being philosophically provocative. He calls his view Heraclitean, in honor of Heraclitus' characterization of time as a war of opposites. The fundamental insight is that "Time is that which allows contradictory things to occur":

> At any *one* time, the statement that a cup is both green and red makes no sense; these are mutually contradictory attributes. At any one time, a single particle can have only one position. However, at different times a particle can have many different positions, as can the cup have many different colours. (Unruh 1988, pp. 254–55; see also p. 2602 of Unruh and Wald 1989)

Time *sets* the values of the other variables, in the sense that at any given time each object takes on exactly one property from any exhaustive and mutually exclusive set (such as position or color), although the property assumed is allowed to vary as the time parameter varies. This is suggestive, but ultimately inadequate. After all, the patterned carpet is allowed to take on different colors in different regions of its spatial extension, just as the colored cup assumes different colors in different parts of its temporal extension. Furthermore, Unruh's proposal is unsatisfactory at the classical level because it depends on a primitive notion of genidentity, which is unlikely to be attractive in the context of field theories.

But he goes on to make a suggestion about how to understand the Heraclitean aspect of time in the context of quantum theories which seems to provide a means to distinguish the spatial from the temporal. Let's suppose that we are given a two dimensional spacetime continuum and a complex function ψ on this continuum which we take to represent the wave function of some quantum particle. Can we distinguish the temporal dimension from the spatial dimensions? Well, let $\{x, y\}$ be an arbitrary set of coordinates. Suppose that we want to calculate the probability of finding the particle in a given region. In order for this to make sense, we will need the total probability to be normalized. But, if ψ is really the wave function, then we expect $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mid \psi(x,y) \mid^2 \ dx\,dy$ to be infinite—at each time, the particle must be somewhere, so the integral over the temporal dimension diverges. The solution is obvious: we should be looking at $\int_{-\infty}^{\infty} \mid \psi(x,y_0) \mid^2 \ dx$ and $\int_{-\infty}^{\infty} \mid \psi(x_0,y) \mid^2 \ dy$, for fixed x_0 and y_0. If we find that the former is constant for all values of y_0, while the latter is extremely badly behaved, then this licenses us to conclude that y is a time variable: integrating over surfaces of constant y gives us normalized probability densities. In this situation, we can view y but not x as setting the conditions for the other variables in the following sense: fixing a value of the time parameter allows us to formulate a quantum theory in which we can interpret the square of the wavefunction as a probability for measurement outcomes for the other variable. This is the Heraclitean role of time in a quantum world.

Of course, it is not straightforward to implement this strategy in the case of quantum gravity. Our discussion above presupposed that we were simply handed the measures dx and dy. But this is to tantamount to knowing the physically relevant inner product on our space of quantum states. And, in fact, to identify the correct inner product is to go a long way towards solving the problem of time in quantum gravity (see especially Kuchař 1993a on this question). Nonetheless, Unruh's suggestion provides a framework in which to talk about the notions of change and time in quantum theories. As such, it has been influential in shaping discussion of the problem of time in quantum gravity, and provides a useful point of departure for our own discussion.

In the next two sections we will sketch some of the most important Parmenidean and Heraclitean approaches to classical and quantum gravity, as well as some of the most telling objections to these proposals. We will begin by sketching a timeless approach to general relativity and quantum gravity in §5, before turning to the details of a couple of Heraclitean approaches in §6.

In what follows, it will be helpful to keep in mind the following picture of the dispute. All parties seem to agree that understanding what the general covariance of general relativity is telling us about change and time is a precondition for the formulation of a theory of quantum gravity. It is true, of course, that in the context of general relativity, we can always cash out talk about time and change in terms Cauchy surfaces in models of general relativity. But, it is maintained, for anyone interested in canonical quantization of general relativity the resources to speak about time and change in quantum gravity must be found in (or imposed upon) the structure intrinsic to the phase space of general relativity. Here one has two options, neither of which is entirely attractive: (a) to embrace the Parmenidean view and attempt to make sense of quantum and classical theories of gravity which are *prima facie* without change or time; or (b) to turn away from gauge invariant interpretations of general relativity, and thus to base one's theories of gravity upon some other interpretation of the significance of the general covariance of the classical theory.

5 Life Without Change

For the Parmenidean, the challenge provided by the general covariance of general relativity is to give an account of the theory in which time and change are not fundamental, but which (i) is consistent with our experience, and (ii) motivates a viable program for quantization. We will begin with (i) before turning to (ii).

5.1 The Classical Theory

Even if it is granted that change is not a fundamental reality, we are nevertheless owed an account of how we can understand the observations of experimental physics and everyday life—observations which would naively seem to involve recording the presence of different properties at different moments of time. Unruh attributes to Bryce DeWitt the suggestion that the accommodation is afforded by time-independent correlations between non-gauge-invariant quantities, a suggestion Unruh himself rejects:

> The problem is that all of our observations must be expressed in terms of the physically measurable quantities of the theory, namely those combinations of the dynamical variables which are independent of time. One cannot try to phrase the problem by saying that one measures gauge dependent variables, and then looks for time independent correlations between them, since the gauge dependent variables are not measurable quantities within the context of the theory.
>
> For example, Bryce DeWitt has stated that one could express measurements in the form of correlations. As an example, one could define an instant of time by the correlation between Bryce DeWitt talking to Bill Unruh in front of a large crowd of people, and some event in the outside world one wished to measure. To do so however, one would have to express the sentence "Bryce DeWitt talking to Bill Unruh in front of a large crowd of people" in terms of physical variables of the theory which is supposed to include Bryce DeWitt, Bill Unruh, and the crowd of people. However, in the type of theory we are interested in here, those physical variables are all time independent, they cannot distinguish between "Bryce DeWitt talking to Bill Unruh in front of a large crowd of people" and "Bryce DeWitt

> and Bill Unruh and the crowd having grown old and died and rotted in their graves." The complete future time development of any set of variables is described in this theory by exactly the same physical variables. The physical variables, those which commute with all the constraints, can distinguish only between different complete spacetimes, not between different places or times within any single spacetime ... The subtle assumption in a statement like the one ascribed to DeWitt, is that the individual parts of the correlation, e.g., DeWitt talking, are measurable, when they are not. (1991, p. 267).

We think that there is a more charitable interpretation to DeWitt's proposal: take it not as a way of trying to smuggle real change through the backdoor but as a way of explaining the illusion of change in a changeless world. The idea is that we measure hyphenated relative observables, such as clock-1-reads-t_1-when-and-where-clock-2-reads-t_2. Such relative observables can be gauge-invariant and hence measurable according to the theory. We then get the illusion of change because we think that we can dehyphenate these hyphenated relative observables and treat each of the component variables as a genuine observable.

Rovelli's proposal for constructing "evolving constants" is the most sophisticated and cogent way of fleshing out this suggestion.[20] In order to avoid the complications of general relativity we illustrate the proposal by means of a toy Newtonian example, which is concocted in such a way to resemble, in relevant features, general relativity as a constrained Hamiltonian system (see the Appendix for a general construction). Consider the Newtonian account of the motion of a free particle on a line. We model this using a Hamiltonian system $(T^*\Re, \omega, H)$ where $T^*\Re = \{(x, p_x)\}$ is the cotangent bundle of the configuration space $\Re$, ω is the canonical symplectic structure, and H is the kinetic energy, $\frac{1}{2}p_x^2$. We now employ the following formal trick, known as *parameterization.* We enlarge the phase space, by adding the time, t, and its canonically conjugate momentum, p_t, and we impose the constraint $0 = p_t + \frac{1}{2}p_x^2 \equiv C$. We also take $H' \equiv C = 0$ as our Hamiltonian. We solve Hamilton's equations to find that our dynamical trajectories $(x(\tau), t(\tau);\, p_x(\tau), p_t(\tau))$ are determined by the equations $\dot{p}_t = 0$, $\dot{p}_x = -\dot{t}\frac{\partial H}{\partial x} = 0$, and $\dot{x} = \dot{t}\frac{\partial H}{\partial p_x} = \dot{t}p_x$, where $\dot{t}$ is arbitrary and the overdot indicates differentiation with respect to the arbitrary time parameter τ. These equations are equivalent to our original equations of motion, p_x =constant and $x = p_x t + x_0$. Thus, each gauge orbit of the parameterized system corresponds to a dynamical trajectory of the original Hamiltonian system.

Since $H' = 0$, we expect that the parameterized system will display some of the same peculiar features as general relativity. The vanishing of the Hamiltonian means that the dynamical trajectories lie in gauge orbits. This means that there are no gauge-invariant quantities which distinguish between two points lying on the same dynamical trajectory. Most strikingly, the position of the particle, x, fails to commute with the constraint, and hence is not gauge-invariant—

[20]See Rovelli 1991a,b,c. For criticism and discussion, see Cosgrove 1996, Hájiček 1991, Hartle 1996, §6.4 of Isham 1993, and §15 of Kuchař 1992.

and so is not measurable under the standard reading of gauge invariance. Thus the parameterized system seems to describe a Parmenidean world in which there is no change—and, in particular, no motion.

This is paradoxical: after all, the parameterized system is empirically equivalent to the original Hamiltonian system, which can be thought of as describing an ordinary Newtonian world. How can we account for this? The obvious response is that we can deparameterize in a preferred way since t, which is supposed to represent the absolute time of Newtonian mechanics, is in principle observable (say, by reading an idealized clock). But suppose we didn't know this, or that we wished to eschew absolute time. How might we describe change, or something enough like change to explain ordinary observations, in the parameterized system?

Choose a global time function on the augmented phase space: a function $T(x, t;\, p_x, p_t)$ whose level surfaces are oblique to the gauge orbits. Consider any phase function F, not necessarily a constant of motion. Define an associated one-parameter family of phase functions $\{F_\tau\}_{\tau \in \Re}$ by the following two requirements:

$$(R_1) \qquad \{F_\tau, C\} = 0$$

$$(R_2) \qquad F_{\tau = T(x,t;\, p_x, p_t)} = F(x, t;\, p_x, p_t).$$

Here $\{,\}$ is the Poisson bracket on the augmented phase space. The first requirement says that each F_τ is constant along the gauge orbits—so that the F_τ, unlike F, are gauge-invariant. In the case at hand, that means that the F_τ are *constants of motion* . The second requirement says that the value of F_τ is equal to the value of the phase function F when the phase point $(x, t;\, p_x, p_t)$ lies on the level surface $T(x, t;\, p_x, p_t) = \tau$. The F_τ are then the evolving constants which can be used to describe change. In our toy example, take $F(x, t;\, p_x, p_t) = x$. F does not commute with the constraint, so it is not gauge-invariant. Take $T(x, t;\, p_x, p_t) = t$. Then using R_2, $F_\tau = x - p - x(t - \tau)$. The value of F_τ on an orbit with initial x_0 is $p_x\tau + x_0$. It is easy to verify that R_1 holds: the F_τ's are gauge-invariant. Being constants of the motion, the F_τ do not change. But the family $\{F_\tau\}$ can be said to "evolve." In our example, the law of evolution is $\frac{dF_\tau}{d\tau} = p_x$. Our tendency to group gauge-invariant quantities into families which can be viewed as "evolving" is supposed to account for our experience of change.

Kuchař (1993a) interprets Rovelli's proposal as saying that change in a non-gauge-invariant quantity can be observed, at least indirectly, by observing the gauge independent quantities F_{τ_1} and F_{τ_2} (say) and then inferring the change Δx in x from $t = \tau_1$ to $t = \tau_2$ to be $F_{\tau_1} - F_{\tau_2} = p_x(\tau_1 - \tau_2)$. Kuchař's objection is that we are not told how to observe τ—we can't do it by observing that the value of t is τ, for t is not an observable in the theory (this echoes the objection of Unruh discussed above).

A possible response is that we *don't* have to observe τ. The F_τ are constants of the motion so it doesn't matter when they are observed—in principle, all the

F_τ could be observed at once. This doesn't make Kuchař happy either: "If all τ is eternally present, all time is irredeemable" (1993a, p. 139). Perhaps another way to put the criticism is to observe that it is hard to see how one would know *which* F_τ one is measuring without measuring τ, which brings us back to Unruh's objection.

These criticisms are misplaced if, as we suggested above, the evolving constants proposal is construed modestly as explaining the illusion of change. Each F_τ is to be taken as a hyphenated relative observable: the-position-x-of-the-particle-when-the-t-clock-reads-τ. We think that there is real change because we (mistakenly) think that we can dehyphenate for various values of the "time" to get differences in the particle position—but the resulting "observable" would fail to be gauge invariant. To be satisfying this line has to be extended to hook up with actual perceptions. Here one might worry that the Kuchař-Unruh challenge comes back to haunt us at the level of neurophysiology if, as the theory seems to demands, all explanations must ultimately be stated in terms of gauge-invariant quantities. But one has to be careful here: we certainly cannot expect the theory to recapitulate our full phenomenology of time—to do so would be to demand that the theory contain a moving now, a demand which physics left behind long ago. But it surely *is* reasonable to demand that there be a place in the theory for models of human beings. In particular, it is reasonable to demand that the theory explain the illusion of change—any gravitational theory which cannot save such basic phenomena as the expansion of the universe will be empirically inadequate.[21]

Kuchař himself seems to admit that the evolving constants framework does meet this latter challenge, and thus *does* provide a way to make sense of time and change in the context of general relativity (see his comments on pp. 138–40 of Ashtekar and Stachel 1991). But even if this is granted, there remain problems with the quantization of general relativity within the evolving constants framework. If these cannot be overcome, this will be a severe blow to the credibility of Parmenidean interpretations of general relativity.

5.2 Quantum Gravity

One of the signal virtues of the Parmenidean approach is that it underwrites an approach to quantizing general relativity which is very clear in its broad outlines (although it is, like every other approach to quantum gravity, extremely difficult in its details). If one regards the general covariance of general relativity as being strictly analogous to the gauge invariance of electromagnetism, then one will treat the quantum constraints of quantum gravity just as one treats the constraints of the quantum theory of the electromagnetic field: one imposes quantum constraints $\hat{h}\psi = 0$ and $\hat{h}_a\psi = 0$—corresponding to the scalar and

[21]One must also face the challenge posed by nonlocality: in (spatially compact) general relativity, each gauge-invariant quantity—and hence each member of some family which we want to view as an evolving constant—is a nonlocal quantity, while we are accustomed to believe that the quantities which we measure are local.

vector constraints of the classical theory—on the space of physical states of quantum gravity.

Heuristically, we proceed as follows. We work in the Schrödinger representation, so that the quantum states, $\psi(q)$, are elements of $L^2(\text{Riem}(\Sigma), \mu)$, and we represent our canonical coordinates, q_{ab} and p_{ab}, via $\hat{q}_{ab}(x)(\psi(q)) = q_{ab}\psi(q)$ and $\hat{p}_{ab}(x)(\psi(q)) = i\frac{\partial}{\partial q_{ab}(x)}\psi(q)$. Then, again heuristically, writing the quantum vector constraint as $\hat{h}_a \equiv \hat{\nabla}_b \hat{p}_a{}^b$, we can show that imposing this constraint amounts to requiring that the quantum wavefunctions be invariant under three dimensional diffeomorphisms. Formally, we can write $\hat{h}\psi = 0$ as

$$\sqrt{\det q}((q_{ab}q_{cd} - \frac{1}{2}q_{ac}q_{bd})\frac{\partial^2}{\partial q_{ac}\partial q_{bd}}\psi[q] - R(q)\psi(q)) = 0,$$

where R is the scalar curvature of q. In this form, the quantum scalar constraint is known as the *Wheeler-DeWitt equation*. One then seeks a representation of an appropriate set of observables on the space of physical states, and looks for an appropriate inner product. Proponents of evolving constants hope to find an appropriate representation of the algebra of classical evolving constants as a set of linear operators on the space of physical states such that: (i) the quantum evolving constants are in fact constants of motion (i.e. they commute with the quantum Hamiltonian); and (ii) there is a unique inner product on the space of physical states which makes the quantum evolving constants self-adjoint (see Ashtekar and Tate 1994 and Ashtekar 1995). Finally, one must construct a quantum Hamiltonian, $\hat{H}$. The classical Hamiltonian can be written as a sum of the classical constraints, so that it is identically zero on the constraint surface which forms the phase space of general relativity. Thus, it is natural to write the quantum Hamiltonian as a sum of the quantum constraints. But since these constraints annihilate the physical states, one concludes that $\hat{H}\psi = 0$.

This quantization program faces some daunting technical problems (see Hájiček 1991, §6.4 of Isham 1991, and §15 of Kuchař 1992 for critical discussion). But there are also conceptual problems. The foremost is, of course, the problem of time: since the quantum Hamiltonian is zero, there is no equation which governs the dynamical evolution of the physical state. Thus there appears to be no change in quantum gravity. Now, we have seen above that the vanishing of the Hamiltonian is a direct consequence of requiring that the quantum constraints should annihilate the physical states. Parmenideans claim that this move is justified by analogy with the successful quantization of other gauge theories. If q and q' are related by a classical gauge transformation, then we expect that $\psi(q) = \psi(q')$. This principle is particularly plausible when q and q' are related by a gauge transformation generated by the vector constraint—demanding that $\hat{h}_a\psi = 0$ is equivalent to demanding that $\psi(q) = \psi(q')$ whenever $q' = d^*q$ for some diffeomorphism $d : \Sigma \to \Sigma$. And this is surely mandatory, since otherwise we could use quantum gravity to distinguish between the (classically) empirically indistinguishable spatial geometries (Σ, q) and (Σ, q').

In the case of the scalar constraint, no such direct geometric justification is available. Here, the Parmenidean must rely upon the general analogy between general relativity and other gauge theories, and upon the following consideration. The scalar constraint of general relativity implements time evolution, just as the constraint imposed on the parameterized Newtonian particle does. Now, we can apply our quantization algorithm to the parameterized particle. The configuration space of the particle is Newtonian spacetime, so the quantum states are wavefunctions on spacetime, subject to the constraint $(\hat{p}_t + \frac{1}{2}\hat{p}_x^2)\psi(x,t) = 0$. In the Schrödinger representation, in which $\hat{p}_t = -i\hbar\frac{\partial}{\partial t}$, the constraint becomes the familiar Schrödinger equation—modulo the fact that the wavefunctions of the parameterized particle are defined on spacetime rather than space.[22]

Thus, the quantum theory of the parameterized particle is intimately related to the quantum theory of the ordinary unparameterized particle. Now, the scalar constraint of general relativity is quadratic in momentum, whereas the constraint of the parameterized particle is linear in p_t, so the Wheeler-DeWitt equation is not even formally a Schrödinger equation—it cannot be solved for the time rate of change of the quantum state. Nonetheless, Parmenideans maintain, one may view the Wheeler-DeWitt equation as encoding all of the information about time and change that is relevant to quantum gravity, in analogy with the quantum constraint of the parameterized particle (since both of them, intuitively, are the quantum versions of classical constraints which generate time evolution).

But, if the Wheeler-DeWitt equation encodes this information, where is the key which will grant us access to it? This is where the evolving constants come in. One presumes that there is, for instance, a quantum evolving constant which corresponds to the classical evolving constant which measures the volume of the universe at different times. By asking for the expectation value of this quantum evolving constant, we can find evolution and change in the *prima facie* changeless world of quantum gravity.

It is at this point that the objections raised by Kuchař and Unruh return with redoubled force. Our discussion of the (classical) evolving constants of parameterized Newtonian particle proceeded smoothly only because we were working with a system in which Newtonian absolute time was merely hidden, and not absent from the beginning. But suppose that we are given a gauge system with a vanishing Hamiltonian, and that this system, like general relativity itself, does *not* arise via parameterization from a Hamiltonian system. Then there will be considerable arbitrariness in the selection of our evolving constants. In particular, we will not have any natural criterion to appeal to in place of R_2 above: we will not know which foliations of our phase space count as foliations by surfaces of constant time, and so our choice of evolving constants will be vastly underdetermined. In particular, we will have no way of guaranteeing that the foliation chosen corresponds to time rather than space—intuitively our family $\{F_\tau\}$ may correspond to the family {the-mass-of-the-object-at-the-point-

[22]Heracliteans will object that this is a crucial caveat: by their lights, the choice of a correct inner product for wavefunctions on spacetime is equivalent to the choice of a Heraclitean time variable. Parmenideans hope to finesse this objection by showing that their approach singles out a unique candidate for the correct inner product for quantum gravity. See Ashtekar and Tate 1994.

x_τ-of-space-at-time-t_0} rather than to {the-mass-of-the-rocket-at-time-$t = \tau$}. And here we are back to the problem discussed in §4: If one doesn't take time and change as fundamental realities, how is one to distinguish between mere spatial variation and true temporal change?

Rovelli himself takes a hard line on this question, and argues that, prior to quantization, any set of evolving constants is as good as any other. One expects, however, that different sets of evolving constants will lead to different quantizations, and that experiment will eventually allow one to determine which sets of evolving constants are viable. Indeed, this situation already arises in the context of quantum mechanics. Hartle (1996) discusses the quantum mechanics of a parameterized Newtonian description which results when non-standard time functions are employed. He finds that predictions that depart from those of standard non-relativistic quantum mechanics can result. This embarrassment can be overcome by restricting to "good" time functions. Of course, it is not evident how such restrictions could be implemented in quantum gravity—so one expects to be faced with a highly ambiguous recipe for quantization. Whether this counts as a strength or a weakness will be a matter of taste.

Here we reach an impasse of a sort which is quite typical of debates concerning the conceptual foundations of classical and quantum gravity. On our reading, the heart of Kuchař and Unruh's objections to the Parmenidean view is to be found at this point. They both possess an intuition which runs directly counter to Rovelli's. They see the distinction between change and variation as fundamental, and doubt that one will be able to formulate a fruitful approach to quantizing general relativity which is blind to this distinction. Thus they see Rovelli's willingness to accept *any* evolving constants as a sign of the conceptual bankruptcy of the Parmenidean approach. Rovelli, of course, rejects this interpretation. For him, tolerance of the radical underdetermination of the evolving constants is part of an attempt to shrug off outmoded classical intuitions about time and change. Indeed, one of the strengths of the Parmenidean approach has been its hints at the discrete structure of quantum spacetime.[23] Both sides agree that the proof will be in the pudding: vindication, if it comes at all, if will come in the guise of a viable theory of quantum gravity. In the mean time, arguments about the proper way forward will continue to be cast in terms of disagreements concerning the nature of change—debates about content and method are inextricably intertwined.

[23] Recent work in Ashtekar's connection-dynamical formulation of general relativity, has produced area and volume operators for quantum gravity which have *discrete* spectra (see Ashtekar 1995 and 1998 for a survey of these results). This is indeed exciting. But the reader is urged to take these results with a grain of salt: (i) the operators in question do not commute with all of the constraints, and hence are *not* observables within the Parmenidean framework; (ii) although these operators have discrete spectra, the family of such operators is parameterized by the family of volumes and areas in Σ, so that an underlying continuum remains. Furthermore, there appears to be some difficulty in defining a physically reasonable version of the quantum scalar constraint for the Ashtekar variables.

6 *Vive le Change!*

Not everyone accepts the Parmenidean approach. Some believe that the analogy between the general covariance of general relativity and principles of gauge invariance of theories like electromagnetism is profoundly misleading. These physicists are skeptical that the Parmenidean approach sketched in the previous section is either mathematically feasible (since they doubt that one will be able to find an appropriate inner product without appealing to a Heraclitean notion of time) or physically meaningful (since they doubt that one would be able to derive sensible physical predictions from a timeless theory). They believe that a Heraclitean time must be found within (or grafted on to) the conceptual structure of general relativity prior to quantization. In this section, we attempt to give the flavor of this approach.

Heracliteans comes in two varieties. They concur that Parmenideans profoundly misunderstand the nature of the general covariance of general relativity, but they disagree as to the correct account. On the one hand, there is a radical wing which forsakes a cornerstone of the traditional reading of general covariance: that in general relativity there is no preferred splitting of spacetime into space and time. On the other hand, there is a more conservative faction which attempts to hew to a traditional understanding of the general covariance of general relativity, while denying that it is a principle of gauge invariance.[24]

There are a number of varieties of radical Heracliteanism (see 6-8 of Kuchař 1992 for an overview). The most straightforward is probably the doctrine that the mean extrinsic curvature is a good time variable for classical and quantum gravity (the mean extrinsic curvature at a point $x \in \Sigma$ of a Cauchy surface with geometry (q, p) is $\tau = q_{ab}p^{ab}/\sqrt{q}$). The point of departure is the observation that there is a large open subset of the space of models of general relativity consisting of spacetimes which admit a unique foliation by surfaces of constant mean curvature (CMC surfaces).[25] If (M, g) is a model which is CMC sliceable, then the mean extrinsic curvature, τ, varies monotonically within the CMC foliation. This observation motivates the following program (see Beig 1994 and Fischer and Moncrief 1996). We restrict attention to that subset of the phase space of general relativity which corresponds to CMC sliceable models. We then solve the vector constraint, by moving to superspace. At this point, we have a gauge system in which the gauge orbits are infinite dimensional. We transform this into a gauge system with one dimensional gauge orbits by stipulating that we are only interested in those points of phase space which represent CMC slices.[26] In effect, we have chosen a foliation for every model, and finessed the

[24]The proposal developed by Unruh and Wald (Unruh 1988 and Unruh and Wald 1989) doesn't quite fit into this classification, since their unimodular time introduces a preferred volume element rather than a preferred slicing. Nonetheless, it is very similar to radical proposals. See §4 of Kuchař 1993b or §4.4 of Isham 1993 for discussion.

[25]The extent of this open set is an open question. See Isenberg and Moncrief 1996 for a recent discussion. Fischer and Moncrief 1996 show that there are three manifold topologies such that if we look at the phase space of CMC sliceable solutions with Cauchy surfaces of such a topology, we find that the reduced phase space is a manifold with *no* singularities.

[26]This is a variety of gauge fixing. In completely fixing the gauge, one kills off gauge freedom

necessity to choose an identification map by working in superspace. The only remnant of the original general covariance of general relativity is the freedom to reparameterize the time parameter, τ. Furthermore, the remaining constraint is linear in the momentum conjugate to τ. Thus, general relativity is now written in the form of a parameterized system: by choosing a distinguished parameterization of τ, one can construct a time-dependent Hamiltonian system whose parameterization is the CMC-reduced form of general relativity. The Hamiltonian, $H(\tau)$, measures the volume of the Cauchy surface of mean extrinsic curvature τ. As noted by Isham (1991, p. 200), it is quite strange to have a theory of the entire universe in which the dynamics is driven by a time-dependent Hamiltonian—usually such Hamiltonians are employed to model the influence of the environment on the system. (This observation also applies to the internal time framework sketched below.)

One hopes that canonical quantization of this Hamiltonian system would lead to a quantum field theory of gravity, complete with a time variable and a (time-dependent) Hamiltonian which governs the evolution of the quantum state via an ordinary Schrödinger equation, and that the expectation value of the volume of the universe, $\langle \hat{H}(\tau) \rangle$, would vary with time. This quantization program has been successfully carried out for $2+1$ general relativity, and is being actively developed in the full $3+1$ case (see Carlip 1998 for the $2+1$ case). One of the remarkable results obtained is that the CMC method of quantization is equivalent to some Parmenidean constructions of $2+1$ quantum gravity. But, of course, all models of $2+1$ general relativity are flat, so that the phase space of the classical theory is finite dimensional, and the quantum theory is a variety of quantum mechanics. One does not expect this sort of equivalence to arise in the $3+1$ case, where the classical phase space is infinite dimensional and the quantum theory is a quantum field theory.

Other radical Heraclitean proposals have similar structures. One much-discussed method is to postulate the existence of a form of matter which allows one to introduce a preferred foliation. For instance, one can postulate the existence of a cloud of dust, each mote of which is a clock. This fixes a reference frame and a time parameter.[27] One then uses this additional structure to reduce general relativity to a parameterized system, which, upon quantization, yields an ordinary Schrödinger equation. Breaking the general covariance of general relativity by introducing preferred frames allows one to introduce a time variable, t, at the classical level which is carried over to quantum gravity. This time

by adding further constraints in such a way that the expanded set of constraints are all second class and the new system is strictly Hamiltonian rather than being gauge; see fn. 8 above. In electromagnetism, one can impose the Lorentz gauge condition, $\partial_a A_a = 0$. This completely fixes the gauge, in the sense that the resulting second class constraint surface intersects each gauge orbit exactly once. The CMC gauge choice only partially fixes the gauge: the resulting constraint surface still has a one dimensional intersection with each of the original gauge orbits.

[27] These will depend upon the Lagrangian which governs the dynamics of the matter: postulating four non-interacting massless scalar fields privileges a system of harmonic coordinates; introducing a cloud of non-rotating and heat conducting dust leads to Gaussian coordinates. See Kuchař 1993b.

variable is Heraclitean: the wave function describing the state of the gravitational field depends on t, and one is able to find an inner product on the space of instantaneous quantum states which is conserved in t. This allows one to make intelligible time-dependent predictions of measurement outcomes.

Such radical approaches view the general covariance of general relativity as an artifact of a particular formulation of the theory. Under this reading, it is true that general relativity can be given a Diff(M)-invariant formulation. But, it is contended, this formulation is by no means the most perspicacious. By using our preferred coordinates to fix the gauge, we can bring to the fore the true physical content of the theory—just as the content of Newtonian physics is most clear when the theory is written in its traditional, non-generally covariant, form. Of course, this reading is vulnerable to the accusation that it betrays the spirit of general relativity:

> foliation fixing prevents one from asking what would happen if one attempted to measure the gravitational degrees of freedom on an arbitrary hypersurface. Such a solution ... amounts to conceding that one can quantize gravity only by giving up general relativity: to say that quantum gravity makes sense only when one fixes the foliation is essentially the same as saying that quantum gravity makes sense only in one coordinate system. (Kuchař 1992, p. 228)

This criticism is extremely telling. To forsake the conventional reading of general covariance as ruling out the existence of preferred coordinate systems is to abandon one of the central tenets of modern physics. Unsurprisingly, radical Heracliteanism has few adherents—such approaches are explored because they are technically tractable, not because they are physically plausible.

Kuchař advocates a more conservative—and ambitious—brand of Heracliteanism.. He articulates a subtle reading of general covariance which differs from both that of the Parmenideans and that of the radical Heracliteans: he denies that general covariance is a principle of gauge invariance, without countenancing the existence of a preferred foliation or a preferred set of coordinates (see Kuchař 1972 for the original proposal, and Kuchař 1992 and 1993a for recent discussions). A good starting point for understanding his approach is to consider the dual role that time plays in a Newtonian world. On the one hand, we can construct a time function, $t(x)$, which assigns a time to each point in Newtonian spacetime. In this guise, time is a scalar function on spacetime. However, we can also think of time as a collection of instants. Because simultaneity is absolute in Newtonian physics, this collection can be thought of as a one-dimensional family, parameterized by t. Equivalently, the real numbers parameterize the ways in which one can embed an instant (surface of simultaneity) into spacetime. Of course, a time function on spacetime suffices to model this role of time as well: the permissible embeddings of instants are just the level surfaces of $t(x)$.

In the context of special relativity one doesn't have a preferred notion of simultaneity, and the two roles of time are no longer so tightly intertwined. One is still often interested in time functions, $t(x)$, on spacetime—especially the time functions associated with inertial observers. But in its guise as the space of instants, time can no longer be thought of as a one-parameter family, since the spirit of special relativity forbids us from identifying the possible embeddings of the instants with the level surfaces of the time function associated with any one inertial observer. In this context there is considerable ambiguity in the notion of an instant. For definiteness, let's fix upon surfaces of simultaneity relative to inertial observers. Then the family of instants will be four dimensional: if we fix a fiducial instant, Σ_0,then an arbitrary instant, Σ, can be reached by applying a time translation and/or Lorentz boost to Σ_0. So we can think of time as being four dimensional in Minkowski spacetime.[28]

Let's now consider a generic model, $\mathcal{M} = (M, g)$, of general relativity. As in the previous cases, it is easy to write down a time function, $t(x)$, on $\mathcal{M}$. One simply requires that its level surfaces be Cauchy surfaces. Let's fix such a time function—and the corresponding foliation of $\mathcal{M}$ by Cauchy surfaces—and enquire after a coherent notion of 'instant' in general relativity. Here, in order to respect the traditional understanding of general covariance, we will want our set of instants to include all Cauchy surfaces of $\mathcal{M}$. Thus, time, *qua* the set of instants, becomes infinite dimensional. We are interested in examining the role that these two notions of time play in the phase space of general relativity.

To this end, we focus our attention on the gauge orbit in the phase space of general relativity which corresponds to $\mathcal{M}$. For a generic $\mathcal{M}$ admitting no symmetries we expect that a given point (q, p) of the phase space represents a three geometry which occurs only once as a Cauchy surface of $\mathcal{M}$. That is, we expect that specifying the tensors q and p on Σ is sufficient to determine a map $X : \Sigma \to M$ which tells us how Σ must be embedded in $\mathcal{M}$ in order to induce the geometry (q, p). Fixing an arbitrary coordinate system on Σ and a coordinate system on M of the form $\{x^\mu\} = \{t, x^a\}$, we find that specifying the twelve independent components of $q_{ab}(x)$ and $p_{ab}(x)$ on Σ determines four real functions on Σ, $\{X^A(x)\} = \{T(x), X^a(x)\}$, which tell us how Σ is embedded in $\mathcal{M}$.

We can think of these maps as functions on the phase space of general relativity: for each point (q, p) of the phase space, $X^A(x)$ is a real number. Following Kuchař we use the notation $X^A(x; q, p]$ to emphasize that each X^A is a function on Σ and a functional (in the physicist's sense) on phase space. This suggests that we could use the X^A and their conjugate momenta, P_B, as coordinates on the phase space, in place of the $q_{ab}(x)$ and $p_{ab}(x)$. Now, of course, knowing the $q_{ab}(x)$ and $p_{ab}(x)$ for a given point of the phase space gives us more information than just the way that the instant is embedded in spacetime—it also tells us about the state of the gravitational field at that instant. Thus, the geometric variables, the $q_{ab}(x)$ and $p_{ab}(x)$, contain information beyond that which is contained in the embedding variables, $X^A(x; p, q]$ and $P_B(x; p, q]$. Indeed, $q_{ab}(x)$ and $p_{ab}(x)$ contain twelve independent components. So specifying

[28]Different notions of 'instant' in Minkowski spacetime produce families of instants of different dimensionalities. See Hájiček 1994 for an analysis and comparison of the distinct varieties of relativistic quantum mechanics which correspond to different notions of instant in Minkowski spacetime.

the geometrical data gives us twelve functions on Σ, whereas specifying the embedding variables gives us only eight. One surmises that there must exist additional variables which represent the true physical degrees of freedom of the gravitational field relative to any given instant (i.e., relative to any fixed values of the embedding variables). Thus, we postulate that the dynamical state of the gravitational field at a given instant is represented by gravitational configuration variables, $\phi^r(x)$ $(r = 1, 2)$, on Σ, together with their momentum variables, $\pi_s(x)$ $(s = 1, 2)$.

So far, we have been restricting our attention to a single gauge orbit of the phase space of general relativity, and depending upon a particular set of coordinates for the corresponding model. More ambitiously, we could look for embedding variables, $X^A(x; q, p]$ and $P_B(x; q, p]$, defined globally on the phase space of general relativity. We then look for a canonical transformation of the phase space of general relativity of the form

$$\{q_{ab}(x), p_{ab}(x) \mapsto \{X^A(x), P_B(x);\, \phi^r(x), \pi_s(x)\}$$

(i.e., we are looking for a change of coordinates which preserves the presymplectic structure). Each of these new canonical variables associates a map from Σ to the real numbers with each point of phase space. We require that the embedding variables satisfy the following two desiderata.

> (I) Global Time. Each gauge orbit of general relativity contains exactly one point corresponding to a given fixed value of the embedding variables.
>
> (II) Spacetime Interpretation. If (q, p) and (q', p') correspond to intersecting Cauchy surfaces of a given model, then we demand that $X^A(x;\, q, p] = X^A(x;\, q', p']$ for points $x \in \Sigma$ which lie in their intersection.

The first condition guarantees that the values assumed by the embedding variables at a given point of phase space do indeed single out a single instant in any given model of general relativity. The second condition guarantees the notion of time as a collection of instants is compatible with the notion of time as represented by a spacetime scalar: the time, $T = X^0(x;\, q, p]$ assigned to a given point x of a relativistic spacetime is the same for all Cauchy surfaces (q, p) passing through that point.[29] If we can find a canonical transformation satisfying these two desiderata, then we proceed to rewrite the constraints in terms of the new coordinates, where they will assume the form $C_A \equiv P_A(x) + h_A(x;\, X, \phi, \pi] = 0$.

[29] It may help to consider how the CMC time fits into this scheme. If we attempt to define $T(x;\, q, p]$ as the extrinsic curvature of (q, p) at x, then we run afoul of the requirement Spacetime Interpretation: if we look at two Cauchy surfaces passing through the same point of spacetime, we expect them to have different extrinsic curvatures at that point. We could, in accord with the CMC proposal discussed above, attempt to define $T(x;\, q, p]$ to be the value of the extrinsic curvature of the CMC slice through x. This satisfies Spacetime Interpretation. But it requires solving the Einstein equations in order to define the time variable on phase space. See §6 of Kuchař 1992 or §4.2.4 of Isham 1993 for discussion.

If all this can be achieved, then we have rewritten general relativity in the *internal* time formulation (time is said to be internal in this formalism because it depends only on phase space variables). It would allow us to reconcile the two roles of time. The internal time on phase space admits an interpretation as a spacetime scalar for any particular model. But general covariance is not broken: there are no preferred foliations or coordinate systems.[30] One can, if one likes, pay special attention to the level surfaces of the time function which the internal time induces on models. But the formalism itself does not privilege these level surfaces: the constraints can be viewed as governing the evolution of the gravitational degrees of freedom between arbitrary instants.

One could go on to apply the Dirac quantization algorithm to the internal time formulation of general relativity. Here the configuration variables are the embedding variables, X^A, and the gravitational variables, ϕ^r. Thus the quantum states will be wavefunctions over the classical configurations of X and ϕ, of the form $\psi(X,\phi)$. We will want to impose the constraints, $\hat{C}_A\psi[X,\phi]=0$. Because the classical constraints are linear in the momentum, the quantum constraints become Schrödinger equations:

$$-i\frac{\partial\psi(X,\phi)}{\partial X^A(x)}=\hat{h}_A(x;\,X,\hat{\phi},\hat{\pi}]\psi(X,\phi),$$

which govern the change in all of the configuration variables under small variations in the embedding variables. Let us denote the space of wavefunctions satisfying these constraints by V_0. At this point, one could proceed as in the Parmenidean program of 5: complete the quantum theory by finding quantum evolving constants, and find an inner product on V_0 which renders them self-adjoint.

Kuchař however, rejects the Parmenidean reading of the significance of the quantum constraints and as a result, he denies that the observables of quantum gravity are self-adjoint operators on V_0 which commute with the constraints. His objections concerning the quantum constraints can be traced back to a subtle difference between his reading of the significance of the general covariance of general relativity, and that of the Parmenideans. As we saw above, his program for quantization takes as its point of departure a formulation of general relativity which fully respects the general covariance of the theory. Kuchař does not, however, subscribe to the Parmenidean dogma that the constraints of general relativity should be understood as the generators of gauge transformations. Rather he draws a sharp distinction between the role of the vector constraint, and that of the scalar constraint. In particular, he holds that the observable quantities of general relativity must commute with the vector constraint, but that they need not commute with the scalar constraint.

The rationale is as follows. In the case of the vector constraint, we can say that "[t]wo metric fields, $q_{ab}(x)$ and $q'_{ab}(x)$, that differ only by the action of

[30] Indeed, Kuchař 1986 argues that the Diff(M)-invariance of general relativity is hidden in the ordinary Hamiltonian formulation of general relativity, but is manifest in the internal time formulation.

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 $x \in \Sigma$."[31] The difference between two geometries, $q_{ab}(x)$ and $q'_{ab}(x)$, related by a transformation generated by the vector constraint is unobservable: it is the difference between identical spatial geometries, which differ only in virtue of *which* point of Σ plays *which* geometrical role. The role of the scalar constraint, h, is very different:

> it generates the dynamical change of the data from one hypersurface to another. The hypersurface itself is not directly observable, just as the points $x \in \Sigma$ are not directly observable. However, the collection of the canonical data $(q_{ab}(x), p_{ab}(x))$ on the first hypersurface is clearly distinguishable from the collection $(q'_{ab}(x), p'_{ab}(x))$ of the evolved data on the second hypersurface. If we could not distinguish between those two sets of data, we would never be able to observe dynamical evolution. (1993a, p. 137)

Or, again, "[t]wo points on the same orbit of [the scalar constraint] are two events in the dynamical evolution of the system. Such events are physically distinguishable rather than being descriptions of the same physical state" (1992, p. 293). Thus, Kuchař believes that there are physically real quantities which do not commute with the scalar constraint of general relativity.

In the internal time formulation, this point will take the following form. $T = X^0(x; q, p]$ has a different status from the $X^a(x; q, p]$ for $a = 1, 2, 3$. The former can be thought of as specifying the instant corresponding to (q, p), while the latter specify how Σ is mapped on to this instant. That is, T specifies a Cauchy surface while the X^a tell us how Σ is mapped on to this Cauchy surface. Thus, the constraint $C_T \equiv P_T(x) + h_T(x; X, \phi, \pi] = 0$ should be thought of as governing time evolution, while the $C_a \equiv P_a(x) + h_a(x; X, \phi, \pi] = 0$ generate gauge transformations which correspond to altering the way that Σ is mapped on to a given Cauchy surface.

Thus, according to Kuchař's analysis, the quantum constraints should not be treated uniformly, as they are within the Parmenidean approach. The quantum constraints $\hat{C}_a\psi = 0$ should be imposed as in the standard approach. This will, as usual, ensure that the theory is indifferent to diffeomorphisms acting on Σ. But the quantum constraint $\hat{C}_T$ demands a different approach. Kuchař recommends the following procedure. Begin by arbitrarily fixing values $T(x) = X^0(x)$ for all $x \in \Sigma$. This specifies an instant in general relativity. Fixing the state of the gravitational field now amounts to fixing the values of $\phi^r(x)$ and $\pi_s(x)$ on Σ, and the classical observables at this instant are just functions of these field variables. It is natural to think of an instantaneous state of the gravitational fields as a wave function $\Psi(\phi^1, \phi^2)$ over field configurations over Σ (with configurations identified if they are related by a diffeomorphism of Σ). An observable will then be any (diffeomorphism-invariant) function of $\hat{\phi}^r(x)$ and

[31] Kuchař 1993a, p. 136. Here and below, we have slightly altered Kuchař's notation to conform to our own.

$\hat{\pi}_s(x)$. In order to bring dynamics into the picture, we impose the constraint $\hat{C}_T\Psi(X,\phi)=0$. This gives us the Schrödinger equation,

$$-i\frac{\partial\Psi(T,\phi)}{\partial T(x)}=\hat{h}_T(x;\,X,\hat{\phi},\hat{\pi}]\Psi(T,\phi),$$

which tells us how the quantum states change under infinitesimal changes of our instant. Now one attempts to find an inner product on the space of instantaneous states which is preserved under the evolution induced by the constraints. If one is successful, then one has a quantum theory of gravity: a Hilbert space, observables, and dynamics. Notice that although the states are gauge-invariant (since they satisfy the quantum constraints), the observables need not be: in general, one expects the expectation values of a function of $\phi^r(x)$ and $\pi_s(x)$ to vary from embedding to embedding (i.e. from instant to instant). Hence, Kuchař's proposal leads to a theory of quantum gravity in which the infinite dimensional internal time plays the role of a Heraclitean time variable.

This quantization procedure has been successfully applied to a number of theories which arise from general relativity by killing infinitely many degrees of freedom (see §6 of Kuchař 1992). Before it can be applied to full general relativity, however, a number of severe technical difficulties must be overcome—including the fact that it appears to be impossible to satisfy Global Time for full general relativity, and the fact that no one has yet been able to write down in closed form an internal time variable which satisfies Spacetime Interpretation.[32] Nonetheless, work continues on the program, in the hope that it is possible to overcome these difficulties (perhaps by modifying the original program).

There are also a number of potential difficulties in interpreting the formalism. The great advantage of the internal time proposal is that it casts quantum gravity into a familiar form: one has a quantum field theory whose states are wave functions over the classical configuration space, and a Hamiltonian which determines the temporal evolution of these states. The chief novelty is that time is now an infinite dimensional parameter, since there are as many ways of specifying an instant as there are Cauchy surfaces in a model.[33] Thus, it seems that the interpretation of such a quantum theory of gravity should be no more (or less) difficult a task than the interpretation of a standard quantum field theory. But this is not quite the case, for three reasons. (i) The fact that the observables are not required to commute with the constraints complicates the measurement problem. If $\hat{O}$ is an observable which does not commute with the constraint $\hat{C}$, then we can find a state, Ψ, of quantum gravity such that $\hat{C}\hat{O}\Psi\neq\hat{O}\hat{C}\Psi=0$. Thus, $\hat{O}\Psi$ is *not* a state of quantum gravity (if it were, it would be annihilated by the constraint $\hat{C}$). So, naively carrying over the formalism of quantum mechanics, it appears that measurement can throw states out of the space of physically possible states. (ii) Since general relativity is a theory

[32]See Schön and Hájiček 1990 and Torre 1992 for the former, and §§1, 2, and 6 of Kuchař 1992 and §§3.4, 4.4, and 4.2 of Isham 1993 for the latter.

[33]But see Hájiček and Isham 1996a,b for formulations of classical and quantum field theories in terms of embedding variables. Also see Torre and Varadarajan 1998 for a problem with unitarity in this context.

of the structure of spacetime, one expects to be able to recover spatiotemporal information from states of quantum gravity, at least approximately. But this appears to be an extremely difficult problem: one expects that the relationship between the geometric data (q, p) and the gravitational degrees of freedom (ϕ, π) is highly nonlocal at the classical level. The inversion of this relationship at the quantum level presents a formidable problem. (iii) One also has to wonder how to make sense of quantum states which are defined as wavefunctions on classical instants—since these instants originally derived their significance from the classical structure of spacetime.

Each of these problems is potentially very serious. By attempting to cast quantum gravity into a familiar quantum field theoretic form, the advocates of internal time may be creating an unintelligible formalism, rather than one whose interpretation is straightforward. Such a turn of affairs would come as no surprise to Parmenideans. On their view, Kuchař's proposal is an attempt to carry classical notions of time over to quantum gravity. From the Parmenidean perspective, it *might* be possible to formulate a consistent theory along these lines, but one shouldn't expect it to be a full theory of *quantum gravity*—since, by all rights, quantum gravity should be a quantum theory of space and time, as well as a quantum field theory of gravity. (As noted above, Parmenideans claim, with some justice, that their discrete-spectra area and volume operators are the first hints of the quantum nature of space and time at the Planck length; see fn. 23 above.) Here we again reach an impasse: Parmenideans and Heracliteans have divergent intuitions about the nature of time and change, and these intuitions condition their taste in approaches to quantizing gravity.

7 The Status of Spacetime

In the preceding sections, we sketched three proposals for quantizing general relativity: evolving constants, CMC gauge fixing, and internal time. These proposals are underwritten by three very different attitudes towards the general covariance of general relativity, and lead to three very different approaches to quantum gravity. Most notably, differences of opinion about general covariance are directly linked to differences of opinion about the existence and nature of change at both the classical and the quantum level. This divergence of opinion cannot be dismissed as *merely* philosophical: it has important ramifications for questions about which quantities are physically real and/or observable in classical and quantum gravity. Indeed, one has every reason to expect that these proposals, if successfully executed, would lead to three inequivalent theories of quantum gravity, which would make very different predictions about the quantum behavior of the gravitational fields.

Before bringing this discussion to a close, we would like to return to the question of the status of the spacetime of general relativity. We proceed by constructing the most plausible interpretation of general relativity which would underwrite each of our quantization procedures.[34] The underlying presumption

[34]We note that among the many proposals for quantum gravity which we have not touched

is that if a given proposal, and no other, were to lead to a successful quantum theory of gravity, that would be a reason to prefer the corresponding interpretation of general relativity over its rivals.

(a) Rovelli's evolving Constants

The motivation for this program is the conviction that general covariance should be understood as a principle of gauge invariance. Thus, one is led to deny that there are any physically real quantities in general relativity which fail to commute with the constraints. As we argued in §3, this drives one towards a relationalist understanding of general relativity—or, perhaps, towards a sophisticated form of substantivalism (more on this below). If spacetime points enjoy existence, then it seems reasonable that a quantity like "the curvature at x" should be physically real. But such quantities do not commute with the constraints, and hence cannot be physically real.[35] Therefore: spacetime points do not exist. Rovelli himself enthusiastically embraces the relationalism which follows from this line of thought.[36]

(b) Constant Mean Curvature as Time

Under this proposal, any admissible model of general relativity comes equipped with a preferred foliation by Cauchy surfaces, as well as a preferred parameterization of the time variable which labels these Cauchy surfaces. The CMC time is absolute in some respects, but not others. There is a preferred notion of simultaneity, and a preferred parameterization of time. But this parameterization is determined by the dynamics of the theory rather than being imposed from outside. Thus, the time which results in this case certainly isn't the absolute time of Newton. Using this time variable, we can write general relativity as a Hamiltonian system whose configuration space is a subset of superspace (the space of equivalence classes of metrics on Σ). Thus, general relativity becomes a theory of the evolution in time of the geometry of space. Here space is best conceived of in relationalist terms: because we take $\text{Riem}(\Sigma)/\text{Diff}(\Sigma)$, rather than $\text{Riem}(\Sigma)$, as our configuration space, we cannot imagine two identical geometries, differently instantiated.

(c) Internal Time

The core of Kuchař's reading of general covariance is that, properly understood, the observables of general relativity should commute with the vector constraint, but not with the scalar constraint—the qualification being essential since, as noted above, Kuchař explicitly allows that the evolving constants

---

upon, there are a number which areparticularly rich in connections to the substantival-relationaldebate. See especially the discussions of Barbour 1994a,b and Smolin 1991.

[35] But recall that the quantities which *do* commute with the constraints are non-local. So today's relationalists find themselves in a somewhat uncomfortable position: there appears to be a mismatch between their relationalist ontology and the non-local ideology which is forced upon them. Traditional substantivalists and relationalists had no such problems—there was a perfect match between their ontology and ideology (spacetime points and the relations between them *or* bodies and the relations between them)

[36] See Rovelli 1997 and his contribution to this volume. See also his 1991d, where he discusses in some detail the relationship between the hole argument and the view that general covariance is a principle of gauge invariance.

proposal provides a coherent framework for understanding observables in the classical theory. The following interpretative stance underlies this approach.[37] In 3, we argued that straightforward substantivalists are committed to the doctrine that each point of the phase space of general relativity represents a distinct physically possible state. This implies that there are physically real quantities which do not commute with any of the constraints of general relativity: presumably, for any two points of phase space which represent distinct physically possible states there must exist a physically real quantity which takes on different values when evaluated at these two points; we could take these quantities to be of the form 'the curvature at point x of spacetime.' Now let (q, p), (q', p'), and (q'', p'') be three points which lie in the same gauge orbit of the phase space of general relativity, and suppose that (q, p) is related to (q', p') by a gauge transformation generated by the vector constraint, and is related to (q'', p'') by a gauge transformation generated by the scalar constraint. Then (q, p) and (q', p') represent the same geometry of Σ—they differ merely as to how this geometry is instantiated by the points of Σ—while (q, p) and (q'', p'') represent distinct geometries. Thus, although (q, p) and (q', p') represent distinct states of affairs for the substantivalists, they represent states of affairs which are empirically indistinguishable. Although there are, according to substantivalists, physically real quantities which distinguish between (q, p) and (q', p'), these quantities are not observables in any literal sense. On the other hand, the states represented by (q, p) and (q'', p'') *are* distinguishable—otherwise we could not observe change. Thus, the physically real quantities which distinguish between (q, p) and (q'', p'') should be empirically accessible. If we now grant that quantum observables correspond to classical quantities which are not only physically real, but also empirically observable (i.e. they do not distinguish between empirically indistinguishable states of affairs), then we see that substantivalists can provide a coherent motivation for the internal time approach.

Do the points of the spacetime of general relativity, then, exist, or not? Given that general relativity is almost certainly false—since it appears to be impossible to marry a quantum account of the other three forces with a classical account of gravity—the only sense that we can make of this question is whether, given our total physics, the best interpretation of general relativity postulates the existence of spacetime. This question will remain open until the nature of quantum gravity is clarified: if distinct interpretations of general relativity mandate distinct quantizations of gravity, then the empirical success of one or another theory of quantum gravity will have repercussions for our understanding of the spacetime of general relativity.

Even at the present stage, however, we can say something about the lessons of the hole argument for our understanding of classical spacetime. (1) Despite widespread skepticism among philosophers, physicists are correct in seeing the hole argument as pointing up a knot of problems concerning: the existence of

[37] We believe that Kuchař himself is committed to this interpretative stance. We do not argue in favor of this claim here. But note the similarity between the reasons adduced below, and the considerations which Kuchař uses to motivate the internal time approach.

spacetime points; the difficult notion of 'observable' in classical and quantum gravity; and the nature of time and change in physical theory.[38] (2) There is no easy solution to the hole argument. We have seen that both traditional relationalism and traditional substantivalism are associated with some extremely difficult technical and conceptual problems when one turns one's attention to quantum gravity—and that these difficulties arise directly out of the doctrines' respective pronouncements on the nature of general covariance. Ultimately, one of the other of these positions may triumph. But it will be because physical and mathematical ingenuity show how the attendant problems can be overcome, not because either position can be ruled out on strictly philosophical grounds internal to general relativity. (3) That being said, we maintain that there is one sort of response to the hole argument which *is* clearly undesirable: the sort of sophisticated substantivalism which mimics relationalism's denial of the Leibniz-Clarke counterfactuals. It would require considerable ingenuity to construct an (intrinsic) gauge-invariant substantivalist interpretation of general relativity. And if one were to accomplish this, one's reward would be to occupy a conceptual space already occupied by relationalism. Meanwhile, one would forego the most exciting aspect of substantivalism: its link to approaches to quantum gravity, such as the internal time approach. To the extent that such links depend upon the traditional substantivalists' commitment to the existence of physically real quantities which do not commute with the constraints, such approaches are clearly unavailable to relationalists. Seen in this light, sophisticated substantivalism, far from being the savior of substantivalism, is in fact a pallid imitation of relationalism, fit only for those substantivalists who are unwilling to let their beliefs about the existence of space and time face the challenges posed by contemporary physics.

We conclude that there is indeed a tight connection between the interpretative questions of classical and quantum gravity. There is a correspondence between interpretations of the general covariance of general relativity and approaches to—and interpretations of—quantum gravity. This correspondence turns upon the general covariance of the classical theory and is mediated via the processes of quantization and the taking of classical limits. One demands that one's interpretation of general relativity should underwrite an approach to quantization which leads to a viable theory of quantum gravity, and that one's understanding of quantum gravity should lead to a way of viewing general relativity as an appropriate classical limit. This provides a cardinal reason to take the interpretative problems surrounding the general covariance of general relativity seriously—at the very least, one wants to know which interpretative approaches to general relativity mandate, open up, or close off which approaches to quantum gravity. More ambitiously, one can hope that clarity concerning the general covariance of the classical theory will provide insights which prove help-

[38]Some physicists might balk at these claim. But, we believe, almost all would agree with the following statement: "the question as to what should be the correct notion of observables in canonical G.R., which is clearly important for any quantum theory of gravity, is not fully understood even on the classical level" (Beig 1994, p. 77). We maintain that once the claim about observables is granted, the others follow.

ful in the quest for a quantum theory of gravity.

A Appendix

In this Appendix we provide a few details about the definitions and constructions mentioned in the text. It falls into four sections, corresponding to material supporting §2.1, 2.2, 2.3, and 5.1, respectively.

A.1 Hamiltonian Systems

Our phase spaces will always be manifolds. These may be either finite dimensional or infinite dimensional. In the latter case, we require that our space be locally homeomorphic to a Banach space rather than to $\Re^n$. For details and for the infinite dimensional versions of the material discussed below, see Choquet-Bruhat et al 1982 or Schmid 1987.

Definition: Nondegenerate Forms. A two form, ω, on a manifold, M, *nondegenerate* if for each $x \in M$ the map $v \in T_xM \mapsto \omega_x(v, \cdot) \in T_x^*M$ is one-to-one. If this map fails to be one-to-one, then there will be non-trivial $v \in T_xM$ with $\omega_x(v, \cdot) = 0$. These are called the *null vectors* of ω.

Definition: Symplectic Form. A *symplectic form* on a manifold, M, is a closed, nondegenerate, two form, ω, on M.

Definition: Hamiltonian Vector Field. The *Hamiltonian vector field*, X_f, of f in (M, ω) is the solution of the equation $\omega(X_f, \cdot) = dH$. When H is the Hamiltonian, we call the integral curves of X_H the *dynamical trajectories* of the system.

Definition: Poisson Brackets. $\{f, g\} = \omega(X_f, X_g) = X_f(g)$.

Construction: Canonical Coordinates. When M is finite dimensional, we can find local coordinates $(q^1, \ldots, q^n; p_1, \ldots, p_n)$ such that ω can be written as $\omega = dq^i \wedge dp_i$.[39] Equivalently, in such coordinates we have:

$$\omega = \begin{vmatrix} 0 & I \\ -I & 0 \end{vmatrix},$$

where I is the $n \times n$ identity matrix. Coordinates of this kind are known as *canonical coordinates*; we speak of the q^i as being *canonically conjugate* to the p_i. In canonical coordinates the equations for our dynamical trajectories assume their familiar form: $\dot{q} = \{q, H\}$ and $\dot{p} = \{p, H\}$. Notice that conservation of energy is a trivial consequence of the formalism: $\{H, H\} = \omega(X_H, X_H) = 0$ since ω is antisymmetric; so H is a *constant of motion* (i.e. H is constant along each dynamical trajectory).

[39] Many infinite dimensional symplectic geometries admit (suitably generalized) canonical coordinates.

Construction: Cotangent Symplectic Structure. Let $(q^1, \ldots, q^n)$ be a set of local coordinates on Q. We then construct a coordinate system on T^*Q of the form $(q^1, \ldots, q^n; p_1, \ldots, p_n)$, where the p_i are just the components of covectors relative to our coordinate system. We can now construct the canonical symplectic form, $\omega = dq^i \wedge dp_i$. That is: the q^i and p_i are canonically conjugate coordinates. This construction is independent of the original coordinate system on Q, $(q^1, \ldots, q^n)$, and can be extended to construct a unique symplectic form for all of T^*Q. Thus the cotangent bundle structure singles out a preferred symplectic structure on T^*Q. This construction can be generalized to the infinite dimensional case.

A.2 Gauge Systems

Definition: Presymplectic Form. A *presymplectic form* on a manifold, N, is a closed two form, σ, with the property that its space of null vectors has the same dimensionality at each point in N.

Definition: Gauge orbit. Two points lie in the same *gauge orbit* iff they can be connected by a curve, all of whose tangent vectors are null vectors of σ.

Construction: Gauge Orbits. The gauge orbits are constructed by integrating the null distribution of σ. That they are manifolds follows from Frobenius' theorem together with the following fact: $[X_f, X_g] = X_{[f,g]}$, where $[\cdot,\cdot]$ is the Lie bracket, so that the map $f \mapsto X_f$ is a Lie algebra homomorphism of $C^\infty(N)$ into $\Xi(N)$, the algebra of vector fields on N. Since the dimensionality of the null space is constant on N, our phase space is foliated by gauge orbits of a fixed dimensionality.

Definition: Dynamical Trajectories. Again, we look at the integral curves of vector fields, X_f, which solve $\sigma(X_f, \cdot) = dH$.

Discussion: Dynamical Trajectories on Constraint Surfaces. If we are thinking of (N, σ, H) as being imbedded 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 $h \in C^\infty(M)$ an *extension* of H to (M, ω) if: (i) $h \mid_N = H$; (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 $X_h \mid_N$ 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 $h \mapsto h + u^a \gamma_a$ 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.

Whereas in the Hamiltonian case Hamilton's equations $\dot{q} = \{q, h\}$ and $\dot{p} = \{p, h\}$ determine a unique dynamical trajectory of the form $(q^i(t); p_i(t))$ through each $x \in M$, we see that in the case of a constrained Hamiltonian system, Hamilton's equations 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 $\dot{q} = \{q, h\} + u^a\{q, \gamma_a\}$ and $\dot{p} = \{p, h\} + u^a\{p, \gamma_a\}$. 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. (Here, for convenience, we have chosen a set of canonical coordinates on (M, ω)).

A.3 Reduced Phase Spaces

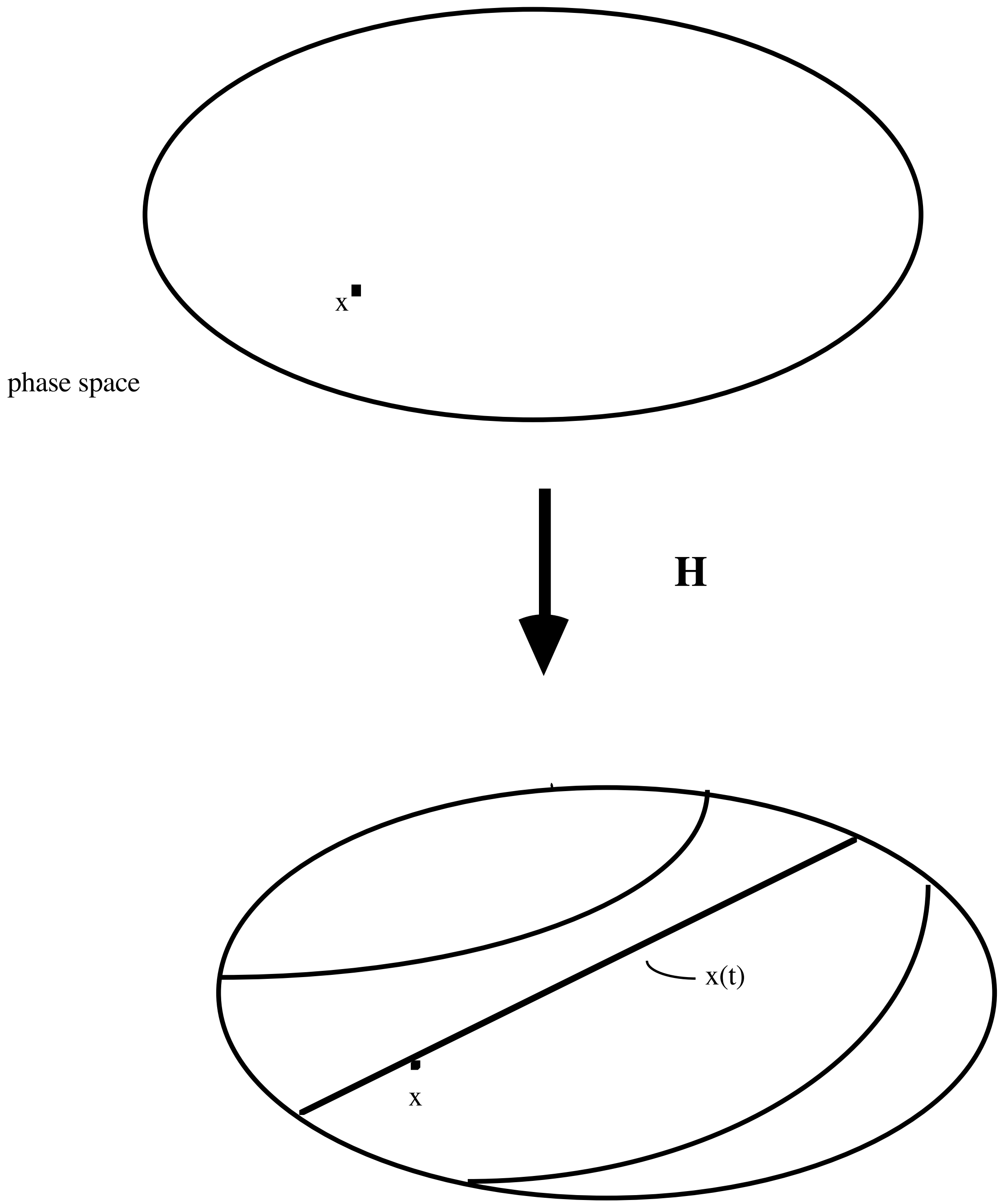
Construction: Reduced Phase Space. The points of the *reduced phase space*, $\tilde{M}$, are the gauge orbits $[x]$ of (N, σ), equipped with the projection topology induced by the projection $\pi : N \to \tilde{M}$. The symplectic form $\tilde{\omega}$ is given by $\pi^*\sigma$. Since H is gauge-invariant, $\tilde{M}$ is well defined by $\tilde{M}([x]) = H(x)$. The set of dynamical trajectories of (N, σ, H) which pass through $x \in N$ projects down to the single dynamical trajectory of $(\tilde{M}, \tilde{\omega}, \tilde{H})$ which passes through $[x]$.

Example: Bad Topology. Here is one way in which this problem can arise. One can construct a constrained Hamiltonian system by starting with a Hamiltonian system (M, ω, H) and imposing the constraint $H = c$. That is, one looks at a surface of constant energy. This is a presymplectic manifold since it has an odd number of dimensions. The gauge orbits of the resulting presymplectic geometry are just the dynamical trajectories of the original Hamiltonian system. Imposing the Hamiltonian $h = 0$ leads to a gauge theory with these gauge orbits as its dynamical trajectories. We can go on to construct the reduced phase space. We simply identify all the points which lie on the same trajectory, and impose the projection topology on the resulting space of dynamical trajectories. What is this reduced phase space like? This depends on the details of the system we started with. If it is integrable, we can find constants of motion (=gauge-invariant quantities) which project down to coordinates on the reduced phase space—the latter will, therefore be a manifold. If, however, our original Hamiltonian system was chaotic, the phase space will be a mess. If our system is ergodic then we will be unable to find constants of motion other than the Hamiltonian, and each trajectory will wander over the entire energy surface. Thus we will be unable to find a sufficient number of gauge-invariant quantities to coordinatize the reduced phase space. Indeed, the topology of the reduced phase space will not even be Hausdorff: since each trajectory of the gauge system approaches every other arbitrarily closely, it will be impossible to separate points of the reduced phase space by open sets.

A.4 Parametrized Systems

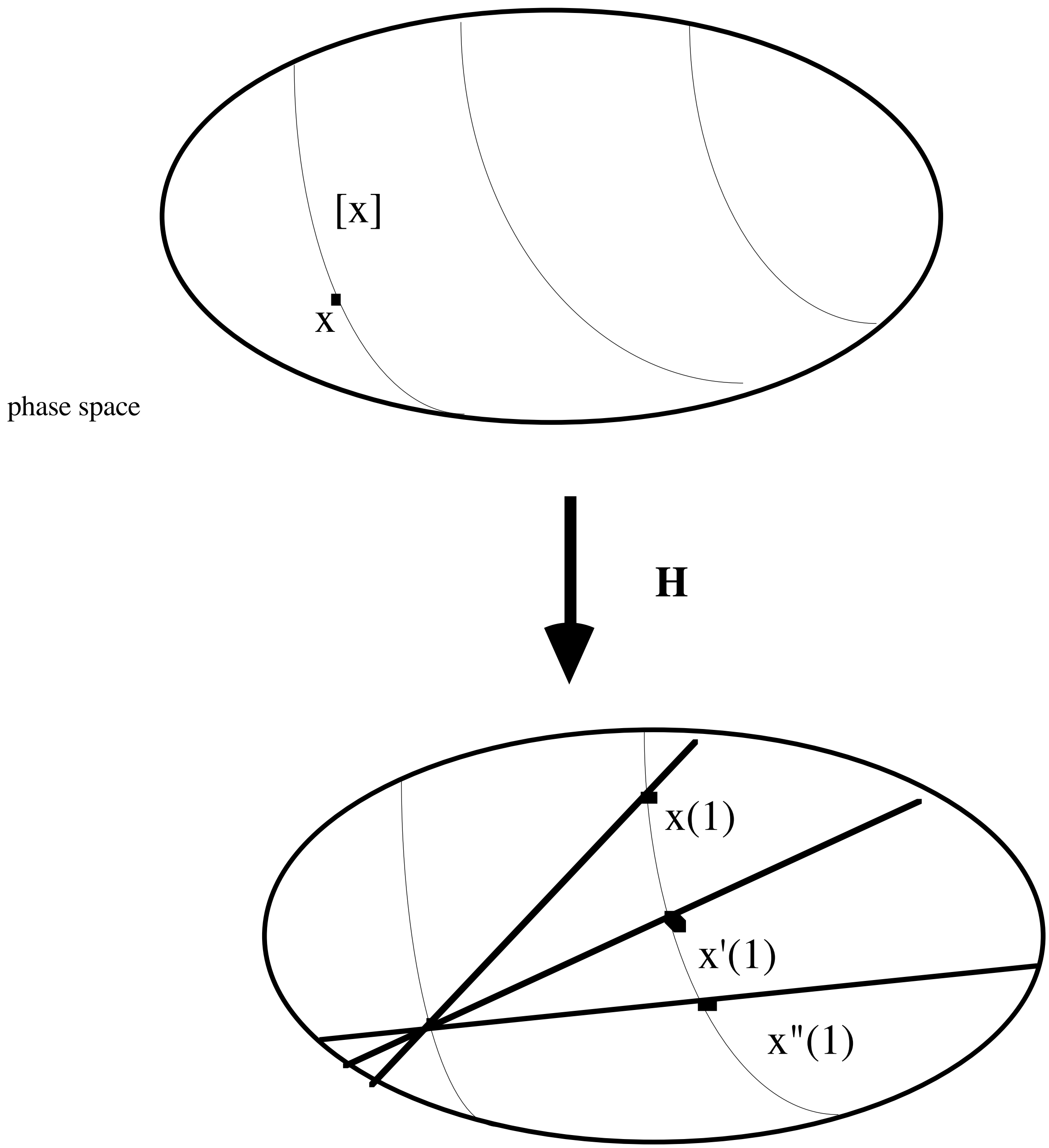
Construction: Parameterization. Let (M, ω, h) be a Hamiltonian system. We construct $M' = \Re^2 \times M$ by adding to M the canonically conjugate variables t and u. Let the symplectic form on M' be given by $\omega' = \omega - du \wedge dt$. Let $H = h + u$ and let N be the submanifold of M' determined by the constraint $H \equiv 0$ (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 $\sigma = \omega' \mid_N$ and the Hamiltonian given by $H = 0$, 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). The gauge orbits are one dimensional and coincide with the dynamical trajectories since $H \equiv 0$ (so that the solutions of X_H 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 $x \in M$, and look at the dynamical trajectory $(q(t), p(t))$ on M. Then the dynamical trajectories through $(t, x) \in N$ will be of the form $(\tau, (q(t), p(t)))$, where $\tau(t)$ is some re-parameterization of time. The gauge orbit in N which corresponds to the trajectory $(q(\tau), p(\tau))$ in M will include all the points in N which are images of the maps $(\tau, (q(t), p(t)))$, for all parameters τ. Thus the loss of the preferred parameterization of time is the price of including time among the canonical variables. The reduced phase space of a parameterized system is just the original Hamiltonian system.

Figure 1: Hamiltonian Systems



46

Figure 2: Gauge Systems



47