UNIVERSITY OF SYDNEY

DOCTORAL THESIS

Symmetry, Ontology and the Problem of Time

On the Interpretation and Quantisation of Canonical Gravity

Karim P.Y. THÉBAULT

Declaration and Copyright

This thesis is an account of research undertaken initially undertaken between August 2008 and August 2011 at the Centre for Time, Department of Philosophy, School of Philosophical and Historical Inquiry, University of Sydney, Australia. Subsequent changes have been in this the final text in line with recommendations of the examiners.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Sections of this thesis have been reproduced in papers for *The British Journal for the Philosophy of Science* (Thébault (2012)), *Symmetry* (Thébault (2011)), and *Foundations of Physics* (Gryb and Thébault (2012)). Where relevant, the copyright rests with the journal publishers. Otherwise it is retained by the author.

Karim P. Y. Thébault
July, 2012

Preface

The true path is along a rope,
not a rope suspended way up in the air,
but rather only just over the ground.
It seems more like a tripwire than a
tightrope.

Franz Kafka

The cause of our confusion is often that the solution to the problem at hand is so near to us that we cannot help but trip over it, and then assume it to be part of the problem itself. So, it will be argued, is the case of the *problem of time in canonical gravity*. Moreover, as we shall see, the realisation essential to understanding the true role of time in canonical gravity is that we have failed to notice that *the solutions* – in the sense of the true dynamical paths – lie in precisely the direction that is perpendicular to where we are accustomed to looking.

The following work is constituted by the confluence of a variety of different ideas, issues and questions that grow out of the interconnected tasks of interpreting and quantising the general theory of relativity. Essentially its focus is on both the classical and quantum facets of the problem of time in canonical gravity and, as such, much attention is given to ideas – both technical and conceptual – targeted directly or indirectly at unpicking the Gordian knot of this multifaceted problem. For all that, as intimated above, in essence our analysis relies on a single, fundamental realisation which we can state in concise but technical terms: the null directions associated with the Hamiltonian constraints of canonical general relativity must be understood as dynamical not unphysical directions. The formal and conceptual basis behind this statement, as well as the relevant qualifications and implications, will be detailed at length during the body of our discussion. Before then it will be useful to give some historical background to our problem.

To understand the history of the problem of time in canonical gravity we must first understand the history of the canonical formulation of gravity itself. Thus, in placing our project in its proper context it is necessary to go back to the physics of the late nineteen fifties and principally the work of Paul Dirac and Peter Bergmann.[1] Dirac's contribution is particularly significant since both the general constrained Hamiltonian formalism (Dirac (1958a)) and the first application of this formalism to general relativity (Dirac (1958b)) can be traced back to him. A few years later Arnowitt, Deser and Misner (Arnowitt *et al.* (1960, 1962)) were able to simplify the formalism and it is their 'ADM' action that we shall study below. The motivation for much of this early work was to find a canonical path towards a quantum theory of gravity. It was expected that if we could gain a good understanding of: i) how to quantise constrained Hamiltonian theories in general terms; and ii) how to formulate general relativity as a constrained Hamiltonian theory, then we would have gone a long way towards the holy grail of a quantum theory of gravity. Unfortunately, things turned out not to be so simple and, despite some admirable progress, half a century later the canonical quantisation program is still beset by severe technical and conceptual problems.

The particular problem that we will be considering in great detail here was identified early on in its classical manifestation and is connected to the fact that the local Hamiltonian functions responsible for generating evolution within the canonical theory are first class constraints. According to the the Dirac prescription (Dirac (1964)), *all* the first class constraints that occur within constrained Hamiltonian theory should be understood as generating infinitesimal transformations that do not change the physical state; thus we have that time is in some sense unphysical! A directly analogous problem can be found within the definition of observables (see Bergmann (1961)). According to Bergmann's definition (which follows from Dirac's prescription) these are represented by phase space functions which have a (weakly) vanishing Poisson bracket with the first class constraints. Since the Hamiltonian is a constraint this means that the observables cannot have any dynamical evolution. Thus, we have two aspects to the classical problem of time; the problem of change and the problem of observables. Below we will consider two strategies which have been developed over the last twenty years and which address these problems by either modifying the Dirac-Bergmann prescriptions or rejecting them outright. Understanding the ontology implied by either of these strategies involves tackling much conceptually challenging terrain and one of the key tasks of this project will be to provide

[1] This is not to discount the notable contributions of others, in particular James Anderson and Arthur Komar. For an early paper which deals with aspects of the problem of time from a non-canonical perspective see Misner (1957). For a fascinating account of the little known 1930s work on constraint theory by Léon Rosenfeld see Salisbury (2007, 2010).

clear guidance as to how the relevant interpretational frameworks should be thought to sit together.

Perhaps more starkly problematic is the quantum mechanical manifestation of the problem of time. In his work on the quantisation of constrained Hamiltonian systems Dirac constructed a technique for canonical quantisation that can be successfully applied to theories such as electromagnetism. Since in that case this *Dirac quantisation* is found to lead to the hugely successful theory that is quantum electrodynamics one would expect that the technique itself has solid mathematical foundations. However, when applied to canonical general relativity Dirac quantisation, which involves the promotion of all first class constraint functions to operators annihilating the wavefunction, leads directly to the Wheeler-de Witt equation (DeWitt (1967)) – '$\hat{H} \mid \psi\rangle = 0$'[2]. This 'wavefunction of the universe' gives at best probability amplitudes on *three* dimensional spatial configurations. Thus, the application of standard quantisation techniques to canonical gravity leads to a quantum formalism that is in a fundamental sense without time.

The question is then: how can we reconcile ourselves to a formalism that represents reality as frozen in an energy eigenstate when our phenomenology abounds with change? Much work over the last few decades has focused upon recovering the impression of dynamics from within the frozen formalism (see Anderson (2010) for a recent review). Here we will develop an entirely different approach whereby it is a problem with the Dirac quantisation technique itself that has led to an *unphysical* timelessness within the conventional canonical quantum gravity formalism. For a class of non-relativistic models we will offer an alternative methodology which involves modifying the Dirac technique such that we are able to retain dynamics. For the full case of general relativity, however, it is not entirely clear how one should proceed towards quantisation whilst retaining dynamics. This is in part due to the various subtleties involved in the two different formulations of canonical general relativity according to the two strategies for solving the classical problem of time mentioned above.

It is in the context of attempting to get a better understanding of the relationship between the problem of quantising gravity and the problem of reconciling two very different formulations of the classical theory, that the lengthly digression into the philosophy of science in the final quarter of this thesis should be seen. There we will examine the extent to which a generic problem of *metaphysical underdetermination* in science might be seen

[2] The inverted commas are because, as we shall see, this expression is far from well defined.

to motivate a structuralist view as to the ontology of physical theory. The particular case that is most relevant to canonical gravity is when this underdetermination is driven by the existence of multiple formulations of a theory, and we shall find suggestive evidence that – given we are dealing with a classical theory – quantisation may be able to give us key insights into what structures are most significant. The final key idea that will be introduced below is that we may be able to invert this idea of quantisation as a guide to structure and use the isolation of common structure between different formulations of a theory as a heuristic towards finding the correct quantisation methodology. Thus, by resolving the underdetermination in the case of our two formulations of classical canonical general relativity, we may be able to gain an invaluable insight into the correct way to proceed towards the quantum theory.

The genesis of this thesis can be traced back to a suggestion made to me by Oliver Pooley during my undergraduate studies at Oxford that it would be interesting to consider the comparison between Julian Barbour's and Carlo Rovelli's *timeless* approaches to quantum gravity as a topic for my fourth year thesis. Whilst working on that earlier project, and during the completion of my masters dissertation at Imperial College London (supervised by Jonathan Halliwell), I spent some time grappling with the various arguments surrounding the treatment of the Hamiltonian constraints of canonical gravity as generators of unphysical transformations. Although I was aware of the Barbour-Kuchař line that such a gauge generating interpretation of the Hamiltonian constraints was incoherent and unfounded, I was still broadly convinced that the mathematical evidence (as well as the 'received' opinion) was against them. During the early stages of my doctoral work at the University of Sydney (supervised by Huw Price and Dean Rickles) I returned to this issue and was prompted to reconsider my earlier opinions by a combination of the then recently posted article Barbour and Foster (2008) and several, much appreciated, discussions with Hans Westman. What, in the end, was for me the key realisation – and what is one of the ideas central to this thesis – is that if we think about the theory of constrained systems in geometric terms then the gauge generating interpretation of the Hamiltonian constraints (within both gravity and non-relativistic models) amounts to moving to a reduced phase space with trivial dynamical structure. Further discussions with Hans, Dean Rickles, Maki Takahashi, Pete Evans and Julian Barbour himself served to cement this key observation in my mind and provided the basis for much of the formal and interpretational material presented below. Moving beyond purely classical theory, I decided to investigate the extent to which this negative conclusion with regard to reduction might be seen to impact upon the conceptual basis of quantum Hamiltonian constraints. The results of this project then led to a collaboration with Sean Gryb (and also initially Tim Koslowski) aimed at finding an alternative method of quantisation which was not predicated upon a problematic reduced space. For non-relativistic models we were able to establish what we believe to be a consistent procedure for achieving a *relational quantisation* that respects the dynamical role of Hamiltonian constraints (the details of this results are included in Chapter 9 below).

To all the above named people I am much grateful for the interactions and discussions that were essential to the completion of this project – and also added greatly to the enjoyment of the entire process! Additional valuable, and again highly enjoyable, interactions have come from the audiences of the numerous talks that I have given on many of the topics covered below – to the the attendees of the weekly Centre for Time graduate meetings, in particular, I am grateful for their patience and philosophical enthusiasm. For help with proof reading large sections of the text below I am grateful to Pete Evans, Dean Rickles, Julian Barbour, Huw Price, Sam Baron and Zahir Thébault. I am also extremely grateful to two of the examiners for their insightful and constructive critical appraisal of the original submitted version of the thesis – many substantial improvements and clarifications have been made based upon their suggestions. I would also like to thank the various friends, family members (in particular my parents) and funding bodies whose support I most certainly would not have been able to do without. Finally, I am much indebted to Gordon Belot, Carlo Rovelli and Thomas Thiemann for their insightful writings on canonical gravity and the problem of time – in particular Belot (2007); Rovelli (2004); Thiemann (2007), these lengthly treatments are notably of great pedagogical as well as technical and philosophical value.

In this vein, and proceeding to a summary of what follows, much of the material within the Part I of this thesis is presented with the aim of providing the reader with a concise review of the substantial amount of mathematics and physics that will prove necessary in our later discussion. An additional aim of Part I is to introduce some of the important connections between the interpretation of classical theories without Hamiltonian constraints (what we shall call *standard* gauge theories) and the quantisation of these theories. Chapter 1 will provide a concise introduction to relevant ideas from differential geometry (§1.1), Lagrangian (§1.2), Hamiltonian (§1.3), symplectic (§1.4) and presymplectic mechanics (§1.5). Chapter 2 will focus upon the philosophical and interpretational stances that can be attached to the various mathematical structures. Chapter 3 principally consists of an introduction to the three quantisation methodologies which will be most significant to our project: geometric quantisation (§3.1), constraint quantisation (§3.2) and path integral quantisation via a Faddeev-Poppov methodology (§3.3). Chapter 4 then presents an interpretative analysis of the relationship between the quantisation techniques, in particular the sense in which they are conceptually, if not formally, predicated upon the viability of classical reduction (§4.1). We then consider the extent to which this facet of quantisation then justifies a *reductionist* type line of the interpretation of the classical theory (§4.2).

The basic structure of Part II is as follows. We begin in Chapter 5 by first giving both a Lagrangian (§5.1) and Hamiltonian (§5.2) description of an extended version of mechanics where time is promoted to a configuration variable. We then proceed to the Jacobi formulation of mechanics within which temporal structure is eliminated altogether (§5.3). The final, and crucial, section of Chapter 5 (§5.4) will then introduce an argument against the applicability of standard gauge techniques (in particular symplectic reduction) to the case in hand. Chapter 6 introduces two non-standard strategies for representing time and observables within non-relativistic reparameterization invariant theory: the emergent time strategy (§6.1) and the correlation strategy (§6.2). Chapter 7 will then consider some of the key interpretational consequences of our discussion in general (§7.1) and of these strategies in particular (§7.2). Chapter 8 will offer some general and introductory ideas about the nature of time in conventional approaches to quantum theory, before we proceed, in Chapter 9, to the introduction of quantum mechanical non-relativistic problem of time. This problem will be seen to have two facets, the first stemming from the reduction issue (§9.1) and the second from the specific exclusion of quantum relational clocks by standard quantisation techniques (§9.2). The nature of these problems is further illustrated by toy model examples (§9.3). We then offer a new proposal for the quantisation of (globally) reparameterisation invariant theories via an intermediary formalism (§9.4) which we will argue to solve the non-relativistic quantum problem of time as we have defined it. We will then consider the structure of the observables of the intermediary formalism in order to demonstrate their ability to be interpreted as representing the *physical* degrees of freedom on the original theory (§9.5), before finally discussing some outstanding interpretive issues implied by the ideas introduced (§9.6.

Part III concerns the full relativistic problem of time. We begin in Chapter 10 with a concise presentation of the canonical formulation of general relativity (§10.1), that is supplemented by an analysis of the relationship with its covariant counterpart (§10.2) and an examination of the role of the Hamiltonian constraints in particular (§10.3). We then proceed to detailing the substance of our first denial of time, first in the context of a motivation taken from standard gauge theory (§11.1) and then in the context of a motivation from reductive spacetime relationalism (§11.2). An argument against the first denial on the basis of dynamical trivialisation will then be presented, together with a rebuttal of the principal line of reasoning that has been employed in its favour (§11.3). Chapter 12 will then present the substance of our second denial on the basis of Machian temporal relationalism (MTR) and the emergent time strategy with which it is assoiacted. After presenting MTR in general terms we will isolate the source of a key problem within its application to canonical general relativity (§12.1). Two possible solutions to this problem will then be evaluated, the first in terms of sophisticated temporal relationalism (§12.2) and the second in terms of a scale invariant formulation of gravity called shape dynamics (§12.3). Chapter 13 will introduce the third denial which is based upon the complete observables scheme that has already been introduced for the non-relativistic case. After a brief restatement of essence of this correlation strategy (§13.1), we proceed to first consider the additional ideas necessary for an application to canonical general relativity (§13.2) and then the philosophical implications with regard to the relative ontological status of space and time (§13.3). Chapter 14 then considers both the implications of the failure of classical reductionism for a Dirac style quantisation of gravity. Chapter 15 consists of some preliminary work towards the application of the ideas of Chapter 9 to the full theory of relativity.

The philosophy of science discussion of Part IV begins, in Chapter 16, with a number of introductory sections. First, we review the two major frameworks for analysing the structure of a physical theory (§16.1). Next, we consider how one of these frameworks may be used to precisely characterise what it is about a physical theory that could be said to be underdetermined (§16.2). Of particular importance will be the specific case within which the underdetermination is driven by multiplicity within the formalisation of a physical theory. We then introduce the position of scientific realism and explain why one might think it to be specifically threatened by underdetermination cases (§16.3). The next section details the various ways our scientific realist may attempt to break the underdetermination by appeal to external criteria (§16.4), before we introduce the alternative position of ontic structural realism (OSR) within which the ontological bite of the underdetermination is supposedly undercut (§16.5). We will also examine both OSR and scientific realism in the context of the historically grounded undermining of ontology that motivated by the argument from *pessimistic meta-induction*, and from this analysis place a set of conditions on an application of OSR being both consistent and substantive. With these conditions in mind, the final section of this chapter (§16.6) will present a scheme for thinking about formulation underdetermination and OSR in the context of quantisation. The following three chapters will then represent case studies for the analysis of the proceeding ideas within three examples of classical formulation underdetermination. Chapter 17 will examine the Lagrangian and Hamiltonian formulations of Newtonian mechanics, and then Chapter 18 will examine the reduced and unreduced formulations of standard gauge theory, before finally, in Chapter 19, we return our discussion to our two rival formulations of canonical gravity. We conclude, in Chapter 20 with a summary of our project together with an analysis of the relevant implications and prospective research avenues that have been illuminated.

List of Key Notational Conventions

Contents

Part I

Standard Gauge Theory

Guide to Part I

In the first part of this thesis we will conduct a review of the classical and quantum mechanical structure of what we call *standard* gauge theories. Although, for simplicity, in our treatment we will assume that we are dealing with finite dimensional systems, the class of standard gauge theories properly includes all gauge theories within which the assumption of an external time parameter is made. Thus, although unfortunately we shall not have space to demonstrate this explicitly, the essential points made within the following discussion are expected to be relevant within both electromagnetism and Yang-Mills theories.[3] We therefore have that at least three of the four fundamental forces of nature can be broadly understood as being described by standard gauge theories.

Importantly, however, many of the details discussed in this part, and the related philosophical conclusions, are not applicable to general relativity or its non-relativistic toy models. Foreshadowing one of the central arguments of this thesis, many aspects of the problem of time will be shown to derive from the misapplication of standard gauge theory techniques to the non-standard gauge theory that is general relativity. However, there is much groundwork that must be done before we get to this crucial point.

Chapter 1 will provide a concise introduction to relevant ideas from differential geometry (§1.1), Lagrangian (§1.2), Hamiltonian (§1.3), symplectic (§1.4) and presymplectic mechanics (§1.5). Chapter 2 will focus upon the philosophical and interpretational stances that can be attached to the various mathematical structures. Chapter 3 consists of an introduction to the three quantisation methodologies which will be most significant to our project: geometric quantisation (§3.1), constraint quantisation (§3.2) and path integral quantisation via a Faddeev-Poppov methodology (§3.3). Chapter 4 then presents an interpretative analysis of the relationship between the quantisation techniques, in particular the sense in which they are conceptually, if not formally, predicated upon the viability of classical reduction (§4.1). We then consider the extent to which this facet of quantisation then justifies a *reductionist* type line of the interpretation of the classical theory (§4.2).

[3] For explicit treatment of these theories (including some important technical subtleties) refer to Sundermeyer (1982); Marsden and Weinstein (1982); Parrinello and Jona-Lasinio (1990). Belot (2007, §5) gives more general details of the geometric structure of standard gauge field theories.

Chapter 1

Geometry, symmetry and constraints

1.1 Elements from differential geometry

Here we give a brief introduction to some of the mathematical concepts essential to our discussion below. For further introductory material on differential geometry the reader is suggested to refer to Baez and Muniain (1994) or Butterfield (2007).

A **Lie group,** G, is a group[4] that is also a differentiable manifold[5] with the property that the product and inverse operations are smooth (i.e., have continuous derivatives). The **action** of a Lie group on a manifold, $\Phi(g, x)$ or $g \cdot x$ where $x \in M$, is a smooth map $\Phi : G \times M \to M$ that implements the identity and associative aspects of the group. The **orbit** of the action through a point on a manifold is a set of points $[x] := \{g \cdot x : g \in G\}$. Under certain conditions the action of G on M is to define a **foliation** of M with the orbits as the **leaves** of the foliation.[6] We can form a set $N = M/G$ known as the **quotient**[7] of M by the group G by considering the set of orbits of the action of G for every point in M i.e., $N := \{ [x] : \forall x \in M\}$.

The simplest Lie group is the additive real group $\mathbb{R}$. It defines an $\mathbb{R}$**-action** $\Phi : \mathbb{R} \times M \to M$ and we associate it with a one parameter group of diffeomorphisms from M to M called a **flow** $\{\alpha^t\}$ through the relation $\alpha^t(x) = \Phi(t, x)$ for $x \in M$. If the flow is well defined for all $t \in \mathbb{R}$ it is *global*, otherwise it is *local*. Every $\mathbb{R}$-action on M induces a unique assignment of a tangent vector, $X \in T_xM$, to every point in M and thus allows us

[4] A set of elements, g, with an identity element, e, within which an operation of combining elements to get another element also in the set (i.e., $g_3 = g_1 \circ g_2$ and $g = g \circ e$) is defined such that it is associative and within which each element has an inverse (i.e., $g \circ g^{-1} = e$)

[5] A space that is locally similar enough to Euclidean space for us to be able to do calculus on it.

[6] See Souriau (1997, p.38 and p.49).

[7] The quotients we will deal with in this thesis will be manifolds, see §1.5 for details. The general conditions for a quotient to be a manifold can be found in Souriau (1997, pp.13-14).

to define a unique **tangent vector field**, $X : x \in M \longmapsto X(x) \in T_x M$. Conversely we can think of a given vector field as generating both an $\mathbb{R}$-action and a flow (the latter we write simply as X^t). Given a vector field on a manifold we can define a family of **integral curves** as smooth maps, $\gamma_x(t) : I \to M$, from a real open interval $I \subset \mathbb{R}$ to $U \subset M$, by considering the local flows (global if $I = \mathbb{R}$ and $U = M$) generated through every point in the manifold (i.e., $\gamma_x(t) : t \longmapsto \Phi_t(x) \; \forall x \in M$). These curves are such that each point in M lies on exactly one such curve and the **parameterisation** of each curve up to a choice of origin is fixed.

Given a vector field, X, on a manifold, M, we can define the **Lie derivative**, $\mathcal{L}_X : f \longmapsto \mathcal{L}_X f$, as an operation on scalar functions, $f : M \to \mathbb{R}$, that gives us the rate of change of f along X.[8] Given another vector field, Y, also on M, we can then consider the commutator between the two relevant Lie derivatives, $\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X \equiv \mathcal{L}_{[X,Y]}$. This defines the vector field $[X, Y]$ which we call the *Lie bracket* of the fields X and Y. The Lie bracket is equivalent to Lie derivative of Y along X and so is also written $\mathcal{L}_X Y$. In effect, it measures the non-commutation of the flows X^t and Y^t. Since the Lie bracket can be understood as constituting the suitable binary operation over a vector space it defines an algebra. This algebra is one instance of a special type of algebra called a **Lie algebra.**

1.2 Lagrangian mechanics

We start with the specification of the set of n independent variables, q_i where $i = 1...n$, which serve to characterise the properties of a mechanical system. These variables are elements of a manifold which we call the **configuration space**, $\mathcal{C}_0$.[9] At a given point $q \in \mathcal{C}_0$ we can define a **tangent space** $T_q\mathcal{C}_0$. The disjoint union of all the tangent spaces of $\mathcal{C}_0$ is called the **tangent bundle** $T\mathcal{C}_0$. The elements of the tangent bundle are pairs $(q, \dot{q})$ of configuration variables q and vectors tangent to those variables $\dot{q}$. For formulations of mechanics with a fixed parameterisation the parameter with which the tangent vectors are defined is unique and may be interpreted as time t (this will prove not to be the case for the theories of mechanics considered in §4). Thus we have $(q, \dot{q}) \in T\mathcal{C}_0$ with $\dot{q} = \frac{\partial q}{\partial t}$.

A curve within the tangent bundle, $\gamma_0 : \mathbb{R} \to T\mathcal{C}_0$, will correspond to a history of a

[8] Explicitly, $(\mathcal{L}_X f)(x) := \frac{d}{dt}|_{t=0} f(X^t(x)) \equiv X(x)f \;\; \forall x \in M$, where $f(X^t(x))$ is the value of f for a given evaluation of the flow generated by X at $x \in M$.

[9] The subscript 0 is used to distinguish the objects introduced here from those of the extended description of mechanics given in §5.

system – a sequence of configurations and velocities. The parameterisation of the curve will be fixed up to a choice of origin and unit by the distinguished time parameter t. This parameter can be taken to vary monotonically along each curve in configuration space. Clearly, for this picture to match up with the physics of the real world we need some restriction on which histories are nomologically possible. This is achieved by defining the **Lagrangian**, $L_0 : T\mathcal{C}_0 \to \mathbb{R}$, and the **action**, $I[\gamma_0] = \int_{\gamma_0} L_0[q_i, \dot{q}_i]dt = \int_{\gamma_0}(T - V)dt$, where T and V are kinetic energy and potential energy respectively. The extremisation of the action, $\delta I[\gamma_0] = 0$, according to the principle of least action leads to the **Euler-Lagrange equations**, $\frac{d}{dt}\left(\frac{\partial L_0}{\partial \dot{q}_i}\right) = \frac{\partial L_0}{\partial q_i}$, that specify a set of parameterised solutions, $\{\gamma_{PS}\} \subset \{\gamma_0\}$, which uniquely determine the physically possible histories of the system given an initial point in $T\mathcal{C}_0$.

1.3 Hamiltonian mechanics

An alternative formulation of mechanics in terms of first order equations is achieved by moving to the **cotangent bundle** of our configuration manifold, the phase space $\Gamma_0 = T^*\mathcal{C}_0$. This is the disjoint union of all the **cotangent spaces** $T^*_q\mathcal{C}_0$ which are themselves defined as spaces of linear functionals on $T_q\mathcal{C}_0$ (i.e., the duals of the tangent spaces). A point in phase space, (q, p), consists of a point in our original configuration space, $q \in \mathcal{C}_0$, paired with a covector at q, $p \in T^*_q\mathcal{C}_0$. These covectors, which we call the conjugate momenta, are given by the Legendre transformation, $\mathcal{FL} : T\mathcal{C}_0 \to T^*\mathcal{C}_0$, which is the map between the configuration-velocity space and the phase space. It can be explicitly constructed using the definition of the **canonical momenta**, $p_i = \frac{\partial L}{\partial \dot{q}_i}$. To fix the dynamics we introduce the Hamiltonian functional, $H_0[q_i, p_i] = p^i q_i - L = T + V$, and derive Hamilton's equations, $\dot{p}_i = -\frac{\partial H_0}{\partial q_i}$ and $\dot{q}_i = \frac{\partial H_0}{\partial p_i}$. The relevant parameterised solutions $\bar{\gamma}_{PS}$ describe the system's dynamics uniquely in the phase space Γ_0 and are isomorphic to the solutions γ_{PS} in the configuration-velocity space $T\mathcal{C}_0$.

1.4 Symplectic mechanics

An elegant and powerful characterisation of mechanical systems is provided by the *symplectic* approach (Abraham and Marsden (1978); Arnold *et al.* (1988); Souriau (1997)). Symplectic is a Greek word first introduced in this context by Weyl (1939). It means roughly 'plaited together' or 'woven'. A symplectic approach to mechanics involves the generalised description of the phase space used above in terms of a natural geometric language with the canonical momenta and configuration variables explicitly represented as woven together.

Above we defined a covector as the dual of a tangent vector, similarly we can define a cotangent vector field or **one-form** as the dual of a tangent vector field. We can generalise these objects to define a k-form as a smooth section of the kth exterior power of the cotangent bundle, $\Omega^k(T^*M)$, of a manifold M. Of particular interest are **two-forms** which are functions $\Omega(x) : T_xM \times T_xM \rightarrow \mathbb{R}$ that assign to each point $x \in M$ a skew-symmetric bilinear form on the tangent space T_xM to M at x (Marsden and Ratiu (1994)). We can transform a k-form into a $k+1$-form by the action of the **exterior derivative**, $\mathbf{d} : \Omega^k(T^*M) \rightarrow \Omega^{k+1}(T^*M)$. It is such that $\mathbf{d}f = df$, $\mathbf{d}(\mathbf{d}\alpha) = 0$ and $\mathbf{d}(f\alpha) = df \wedge \alpha + f\mathbf{d}\alpha$ where α is a k-form and df is the differential of f. Here we have introduced the totally anti-symmetric **wedge product**, $\wedge$, which for a pair of one forms $\theta, \phi \in \Omega^1(T^*M)$ can be simply expressed in terms of the usual tensor product as $\phi \wedge \theta = \phi \otimes \theta - \theta \otimes \phi$.

Given a general cotangent bundle, T^*M, we can always define a corresponding **Poincaré one-form**[10], θ, in terms of a sum of wedge products between a covector and the total differential of the vector it is paired with. Thus for our phase space, Γ_0, the Poincaré one form is $\theta = p_i dq^i$. If we then take the exterior derivative we get a two-form:

$$\Omega_0 = \mathbf{d}\theta = \mathbf{d}(p_i dq^i) = dp_i \wedge dq^i \tag{1.1}$$

This two-form is called a **symplectic two-form** and is both closed ($\mathbf{d}\Omega_0 = 0$) and non-degenerate (if $\Omega_0(X_f, X_g) = 0$ for all $X_f \in TM$ then $X_g = 0$). A manifold endowed with a symplectic two-form constitutes a **symplectic geometry** (M, Ω_0). Significantly, if we are given a smooth function, f, on a manifold endowed with a symplectic two-form then we immediately define uniquely a smooth tangent vector field X_f through the map $f \longmapsto X_f$ given to us by $\Omega_0(X_f, \cdot) = \mathbf{d}f$. The uniqueness of the vector field is guaranteed by the non-degeneracy of Ω_0.

The relation between symplectic geometry and the Hamiltonian theory of mechanics

[10] See Westenholz (1978, pp. 392-4) for more details.

outlined above can be seen immediately since Hamilton's equations can be written:

$$(\dot{q}_1, ..., \dot{q}_n, \dot{p}_1, ...\dot{p}_n) \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \left(\frac{\partial H_0}{\partial q_1}, ..., \frac{\partial H_0}{\partial q_n}, \frac{\partial H_0}{\partial p_1}, ..., \frac{\partial H_0}{\partial p_n} \right) \tag{1.2}$$

where I is the $n \times n$ identity matrix. This expression is an unknown vector multiplied by a matrix and set equal to known vector. It is equivalent to

$$\Omega_0(X_{H_0}, \cdot) = \mathbf{d}H_0 \tag{1.3}$$

which is an unknown tangent vector field (the Hamiltonian vector field X_H) contracted with a two-form and set equal to the exterior derivative of the Hamiltonian, H. Thus we can see Hamilton's equations have an immediate connection with symplectic geometry. The dynamics of a system can be totally specified by the triple $(\Gamma_0, \Omega_0, H_0)$, where Γ_0 is our phase space manifold (cotangent bundle), Ω_0 is the symplectic two form, and H_0 is the Hamiltonian function on Γ_0. Together these three elements fix the value of the Hamiltonian vector field, X_{H_0}. It is the integral curves of this vector field that correspond to the parameterised phase space solutions $\bar{\gamma}_{PS}$ that we associated with the physical histories above.

The Hamiltonian vector field that we have just defined is unique. This implies that it will generate a unique $\mathbb{R}$-action on phase space. This Hamiltonian $\mathbb{R}$-action, and the associated Hamiltonian flow[11], are what we conventionally identify as temporal evolution since they take us from a point in phase space (instantaneous state of a physical system) to a second point (state) that is t units along a solution (physical history). Thus, we see that there is a intimate connection between the Hamiltonian and time.

This connection is made even more explicit by the introduction of the **Poisson bracket**, which is a special case of the Lie bracket, that can be defined via the symplectic two-form for any pair of functions, $f, g \in C^\infty(\Gamma_0)$, as $\{f, g\} := \Omega_0(X_f, X_g)$. The Poisson bracket can be related to the action of a vector field on a smooth function $\{f, g\} = X_g(f) \equiv df(X_g) \equiv \mathcal{L}_{X_g}(f)$. This means that if we take the Poisson bracket of the Hamiltonian with an arbitrary smooth function we will get the change of this function along the flow

[11]It is important to note here that the word flow will always be used in the precise mathematical sense given in §1.1 and has only a tenuous relationship with the (arguably ill-defined) metaphysical notion that goes under the same name. See Price (2009) for discussion of the problems with the metaphysical notion of flow.

defined by the Hamiltonian vector field. This is equal to the variation of the function with respect to the flow parameter of X_{H_0} which is, of course, how change with respect to time is represented:

$$\{f, H_0\} = X_{H_0}(f) = \frac{df}{dt} = \dot{f} \tag{1.4}$$

Conversely, the commutation condition $\{f, H_0\} = 0$ indicates that a function is conserved – it does not change with respect to time.

1.5 Presymplectic geometry and symplectic reduction

A physical system within which a Lie group, G, acts on the tangent bundle, $T\mathcal{C}_0$, such that the Lagrangian, L, is invariant and the group is local (i.e., it can be parameterised in a natural way by a family of arbitrary functions on space-time) is said to display a gauge symmetry. In such systems the assumption that the Legendre transformation is an isomorphism which was implicit in our construction of mechanics above no longer holds. This is because the bijectivity of the map $\mathcal{FL} : T\mathcal{C}_0 \to T^*\mathcal{C}_0$ is dependent on the Lagrangian being such that it determines tangent vectors $\dot{q}$ uniquely through the definition of the canonical momenta. Gauge symmetries $g \in G$ manifestly subvert this since we have that $L(q', \dot{q}') = L(gq, g\dot{q}) = L(q, \dot{q})$ for $\forall g \in G$. In phase space terms the existence of a gauge symmetry group corresponds to the p_i's and q_i's not all being independent - there exists some functional relationship between them of the form $\phi(p, q) = 0$. We call such functions **constraints**. These particular constraints are often called *primary* constraints following the terminology introduced by Dirac and Bergmann. Primary constraints are distinguished as being those resultant directly from the fact that the conjugate momenta are not independent functions of the velocities, rather than the *secondary* constraints which result from the application of consistency conditions that ensure the primary constraints are conserved. Such subtitles will not will not be important to our purpose. The reader is referred to the classic discussion of Dirac (1964) for more details.

Geometrically we can understand the collection of all the constraints, ϕ_j where $j = 1, ...m$, as defining an $(2n - m)$-dimensional sub-manifold, $\Sigma = \{(p, q) \in \Gamma_0 | \forall_j : \phi_j(p, q) = 0\}$, within phase space, Γ_0, that we call the constraint surface. Given the definition of this surface we can the make the crucial distinction between *first class constraints*, which have a vanishing Poisson bracket with all the other constraints when restricted to the constraint surface, and *second class constraints*, which do not. Here we will assume that all constraints are first class – this is justified by the fact that we have explicitly assumed their origin to be within a local symmetry group.

The phase space itself will, as in the unconstrained case, have a symplectic geometry characterised by the pair (Γ_0, Ω) – where Ω is again a closed and non-degenerate two-form constructed by taking the total differential of the Poincaré one form $\theta = p_i dq^i$. However, points in this space which do not lie on the constraint surface will not correspond to physically possible states since they constitute solutions which violate the gauge symmetry. These points are *inaccessible* or *merely unphysical* in the language of Rickles (2008, p.177) and their identification as representing physical states would represent a violation of the law of nature that the gauge symmetry encodes. It is the geometry particular to the class of points lying on the constraint surface that is nomologically significant.

We can characterise the geometry of the constraint surface explicitly by first restricting θ to Σ to get a new characteristic one form, $\tilde{\theta} = \theta_{|\Sigma}$. The total derivative of $\tilde{\theta}$ will then give us a two-form $\omega = \mathbf{d}\tilde{\theta}$ which endows the constraint manifold with the geometry (Σ, ω). This new two-form will be closed but whether it is degenerate or not depends on the particular properties of the constraint surface itself. In cases where it is non-degenerate we again have a symplectic geometry and the dynamics is as described above only now with the triple $(\Sigma, \omega, \tilde{H}_0)$ defining the system (where $\tilde{H}_0 : \Sigma \to \mathbb{R}$ is the restriction of H_0 to Σ).

In the case that ω is degenerate, however, we have a **presymplectic** geometry and our regular description of dynamics is no longer available to us. This is because presymplectic geometries have a degenerate structure that does not allow us to associate a unique vector field with every smooth function. This means that we are not provided with a straightforward characterisation of time evolution either via a unique $\mathbb{R}$-action or by the usual Poisson bracket with the Hamiltonian. Even more worryingly, the existence of local symmetry groups allows for indeterministic or *underdetermined*[12] evolution since at a given point the degeneracy of the Hamiltonian vector field allows for multiple mathematically distinct but dynamically equivalent solutions irrespective of the path leading up to that point. Thus, it would seem that the degeneracy inherent in presymplectic geometries is of a pernicious variety such that we can no longer establish a direct representational rela-

[12]This sense of underdetermined should not be confused with the metaphysical notion of underdetermination that will be considered in part IV. Whereas in that case we have the possibility of different kinds of ontologies, here we rather have the possibility of different future ontologies (of the same kind) given a shared past ontology.

tionship between the relevant mathematical and ontological objects – there is no longer a one-to-one correspondence between the phase space solutions and the physical histories which are distinguished by unique values of the action and so our theory is underdetermined.

To get a better hold on the nature of this degeneracy we can define the null tangent vector space $N_x \subset T_x\Sigma$ as the collection of vectors that satisfy the equation $\omega(X, \cdot) = 0$. This is equivalent to the null space or kernel, $Ker(\omega)$, of the presymplectic form. A kernel of dimension greater than zero is characteristic of the non-trivial structure of the presymplectic form just as a kernel of dimension equal to zero is characteristic of the trivial structure of the symplectic form. An equivalence relation between two points $x, y \in \Sigma$ can be defined based upon the condition of being joined by a curve, $\bar{\gamma} : \mathbb{R} \to \Sigma$, with null tangent vectors. Sets of points for which this equivalence relation holds are submanifolds called gauge orbits, $[x]$, and we say that the action of our presymplectic form is to partition phase space into these orbits. Equivalently we can say that the orbits are defined by the integral curves of the null vector fields of ω. The non-uniqueness that we understood in terms of the existence of gauge orbits is, therefore, also characterised by $Ker(\omega)$.

Critically for our purposes the quotient $\Pi_R = \Sigma/Ker(\omega)$ will necessarily be both symplectic and a manifold. The first is assured since the quotient is with respect to a *sectional foliation*.[13] The second is assured because the quotient is of a presymplectic manifold with respect to the kernel of its own presymplectic form and it can be shown that this implies that the resulting quotient manifold will be endowed with a closed two-form with a kernel of zero dimension – i.e., it will have a symplectic geometry.[14] We can now represent evolution in terms of a unique $\mathbb{R}$-action defined in Π_R. We call Π_R the reduced phase space and using the projection map $\pi : \Sigma \to \Pi_R$ can define the symplectic geometry (Π_R, Ω_R, H_R) where Ω_R is the two-form whose pullback to Σ by π is ω (i.e., $\omega = \pi^*\Omega_R$ where $\pi^* : \Pi_R \to \Sigma$). An equation of the form $\Omega_R(X_{H_R}, \cdot) = \mathbf{d}H_R$ then gives us a unique Hamiltonian vector field along with the associated Poisson bracket and $\mathbb{R}$-action that allows us to uniquely represent both time and the physical histories uniquely within our formalism.

The pullback by π also allows us to consider the properties that smooth functions on

[13]See Souriau (1997, p.42 and pp. 82-3). It is a sectional foliation because the orbits which partition Σ constitute manifolds which are suitably transverse.

[14]See Souriau (1997, theorem 9.10).

the reduced phase space will have with respect to the constraint manifold. Given such a function, $f_R \in C^\infty(\Pi_R)$, we can define $f_\Sigma \in C^\infty(\Sigma)$, by $f_\Sigma = \pi^* f_R$. Since points connected by a gauge orbit on Σ will be represented by a single point on Π_R we have that f_Σ will be constant along such gauge orbits. We can also talk about functions on the full phase space as being constant along gauge orbits. Since the constraints are by definition functions of the form $\phi_j : \Gamma_0 \to \mathbb{R}$ the symplectic form on phase space will associate them each with a vector field X_{ϕ_j}. If we then take the Poisson bracket between them and an arbitrary function, $f \in C^\infty(\Gamma_0)$, we will have $\{f, \phi_j\} = \Omega(X_f, X_{\phi_j})$. On the constraint surface it must be the case that the X_{ϕ_j} coincide with the null vector fields N – the integral curves of which are the gauge orbits. So, given that on the constraint surface f must be a function which is unchanging along the gauge orbits, the definition of the Poisson bracket implies that the expression $\{f, \phi_j\}$ must vanish on the constraint surface – i.e., we have that $\{f, \phi_j\} \approx 0$, where the weak equality is understood to mean vanishing upon the constraint surface.

We can therefore distinguish a class of functions on phase space, *Dirac-Bergmann observables*, by the satisfaction of three equivalent conditions:

(i) Constancy along gauge orbits on the constraint manifold

(ii) Weakly commuting with all the constraints

(iii) Equivalence to a function on the reduced phase space

The name observable seems sensible since it is only these functions that are specified uniquely for every value of the flow parameter defined by the vector field generated by the reduced Hamiltonian, H_R. Thus, given our reliance on an underlying symplectic structure to define time, precise restrictions are placed upon the mathematical objects with which we would want to associate physical quantities.

This idea of passing from a presymplectic to a symplectic manifold by quotienting with respect to the kernel of the presymplectic form is what we will call symplectic reduction and has an important connection[15] with Dirac's theory of constraints. In particular, in cases (such as those considered in the next section) where there is only one primary constraint (and no secondary constraints) the application of symplectic reduction is identical to following the Dirac procedure in that it leads to the same conditions on observable

[15] See Gotay *et al.* (1978); Pons *et al.* (1999) for explicit examination of this connection.

functions we have just outlined. A theory in which all the primary constraints that are first class (i.e., have a weakly vanishing Poisson bracket with all the other constraints) are gauge generating is said to obey Dirac's theorem (Barbour and Foster (2008)) and we can therefore say that the applicability of symplectic reduction is equivalent to satisfaction of Dirac's theorem in all theories with a single primary constraint. There is also a very close connection between symplectic reduction and the Dirac interpretation of first class constraints for theories with multiple primary and secondary first class constraints. Although, in such cases the formal definition of the reduced phase space is much more subtle (see Gotay *et al.* (1978) for the full procedure), we may still think of the reduced phase space as a space of gauge orbits constructed via something analogous to the quotienting out the null directions associated with the first class constraints.

Chapter 2

Reductionism, gauge symmetry and histories Haecceitism

The identification between gauge theories treated according to Dirac's constraint procedure and the re-construction of such theories in terms of reduced phase spaces arrived at via symplectic reduction has important interpretational consequences. As we have seen above conventional Hamiltonian mechanics can be characterised in terms of a phase space which has a symplectic geometry and within which solutions (the integral curves of the Hamiltonian vector field) are in one-to-one correspondence with physical histories. In these circumstances it seems natural to identify the phase space as a *possibility space* since each point can be considered to represent a distinct possible instantaneous physical state and each curve a distinct possible physical history. On the other hand, when we have a constrained Hamiltonian system the relevant phase space is clearly not a suitable candidate for a possibility space it contains *inaccessible* points (i.e., those not on the constraint surface) which can not be thought of as representing physically possible states. Furthermore, even if we exclude such points and focus on the *physical* section of phase space (i.e., consider only points on the constraint surface) then we again do not have a natural candidate for a possibility space since the weaker presymplectic geometry only equips us with an equivalence class of solutions corresponding to each physical history. This leaves the theory open to pernicious underdetermination which is such that if points are identified as representing distinct instantaneous states, then specifying an initial sequences of states fails to uniquely determine future states.

Given a standard gauge theory, such as electromagnetism, which is manifestly deterministic in the sense of giving unique predictions for all measurable quantities, a *literal* interpretation of the physical phase space as constituting a possibly space would then naturally lead to the conclusion that there were physically real quantities that are not measurable. For the case of electromagnetism this would mean that even though the theory only allows for the values of the electric and magnetic fields to be measured, one should still interpret the value of the vector potential seriously as a physical magnitude. As pointed out by Belot and Earman (2001, p.222) such an approach seems rather strange and would require us to construct a highly unorthodox account of the concept of measurement. However, such a version of physical phase space literalism still constitutes a consistent option and we must investigate the viability of the other options before (potentially) dismissing it.

An alternative would be to retain phase space literalism, but to deny that the difference between phase space points should be *cashed out* in terms of real but unmeasurable quantities. Rather, we can take the difference between instantaneous states (as represented by gauge related points in phase space) to be grounded in terms of the differing roles that these states play in the context of the histories (i.e., phase space curves) of which they are part.

Following Lewis (1983), we can designate as *Haecceitists* those who admit 'nonqualitative determinants of cross-identification' (p.19) between entities or objects in distinct worlds or structures. To adopt such a position is to allow for real differences which are only with respect to which objects play which role within the structure; since one is allowed to cross-identify each of a pair of qualitatively identical objects whose roles are permuted between two structures, we may ground a non-qualitative differentiation *of the structures* in terms of the cross-identification *of the objects*.

A literal way of interpreting a possibility space – i.e., each point represents a distinct instantaneous state – can then be understood in terms of a *histories Hacceitism* position that does not include real but unmeasurable quantities. We can seen this since: i) The literal interpretation involves us considering as distinct two histories represented by sequences of points which differ solely with respect to a gauge transformation; ii) Such a difference is only with regard to which instantaneous states (represented by points) play which roles; iii) This means that if we take a history to be the relevant structure and instantaneous states (labelled by the points to which they correspond) to be the relevant objects, then the ontological difference between gauge related histories in the literal interpretation can be naturally *cashed out* in terms of our notion of histories Hacceitism.

(This is not to imply that there many not be other methodologies to ground such differences. For example Butterfield's (Butterfield (1988)) response to the hole argument in general relativity makes use of counterpart theory rather than histories Haecceitism to establish a non-qualitative yet ontologically significant difference between gauge related histories. One might also, of course, seek to ground Hacceitistic differences in terms of objects within the relevant instantaneous states or histories. This more conventional, more theory specific approach will be introduced in the specific context of space and time within Chapter 7.)

Now, although histories Haecceitism clearly does not, by definition, allow for indeterminism with regard to (unmeasureable) qualitative quantities *à la* our original version of physical phase space literalism, it does, again by definition, allow for a species of ontological indeterminism: Two sequences of instantaneous states can initially coincide but then differ in a real but non-qualitative manner as determined by a purely haecceitistic differentiation between the histories. Given we are dealing with classical, deterministic theories, one might wish to construct a position such that what we treat as the ontology is entirely deterministic, and histories Haecceitism clearly will not allow this.

An anti-Haecceitist will deny the possibility of non-qualitative determinants of cross-identification and so will disavow exactly the haecceitistic differentiation that allows for two gauge related sequences of points in a possibility space to represent distinct structures.[16] Thus, by adopting *histories anti-Haecceitism* we can relieve ourselves of the burden of having to endorse ontological indeterminism by instituting a many-to-one relationship between gauge related sequences of points on the constraint surface and the unique sequences of instantaneous states they represent.

We thus have three possible interpretations of the physical phase space of a standard gauge theory. The first leads us to allow for real but non measurable quantities, and constitutes a literal, qualitatively ontologically indeterministic interpretation of physical phase space. The second (histories Haecceitism) leads us to allow for real difference with regard to which instantaneous states play which roles within histories, and constitutes a literal, non-qualitatively ontologically indeterministic interpretation of physical phase space. The third (histories anti-Haecceitism) leads us to deny difference with regard to which instantaneous states play which roles within histories, and constitutes a non-literal, entirely ontologically deterministic interpretation of physical phase space. The third options seems to my mine the most attractive (not least on grounds of ontological parsimony), however there is not in principle reason to exclude any of them.

Although providing space for an attractive interpretation of the possibility space structure found in gauge theory the anti-Haecceitist approach does nothing about removing

[16]They need not, however, also deny primitive identity of the objects concerned (i.e., instantaneous states) since such primitive identity may be conceived of contextually. See Ladyman (2007) on this point.

what would seem like superfluous mathematical structure – to dispense with this surplus structure we need to move to the reduced phase space. Now, this space has obvious interpretational benefits since, as seen above, if all goes well the reduced space will be a symplectic manifold with the integral curves of the reduced Hamiltonian vector field naturally identified as representing physical histories and points as representing physically distinct instantaneous states. Thus the reduced space will, by definition, not feature any underdetermination associated with gauge symmetry and if we endow it with the privileged status as our fundamental possibility arena we reap the reward of recovering the ability to use our conventional representational scheme for theories which display gauge symmetry. Since we have regained a one-to-one correspondence between possibility space points and physically distinguishable instantaneous states the applications of notions, such as histories Haecceitism/anti-Haecceitism discussed above, becomes unnecessary. The superiority of, when possible, reduction as an interpretational stance has been advocated principally by Gordon Belot and John Earman (Belot (2000, 2003); Earman (2003); Belot and Earman (1999, 2001)). We will call it the *reductionism* with regard to constrained Hamiltonian theory and a close association can be made between it and Dirac's theorem as defined above – in fact, it would seem fair to say that the reductive philosophical stance is the natural interpretational consequence of a strict reading of Dirac's theorem. In the next chapter we shall consider three methodologies for the quantisation of standard gauge theories. The connection between these and reductionism will then be considered in §4.2.

A seeming alternative to reductionism that also leads to a deterministic yet literal phase space formalism is to apply a *gauge fixing* such that we consider as our physical phase space a manifold defined by the constraint surface and second manifold of the same dimensionality. The gauge fixing is picked such that the intersection surface is a submanifold within which exactly one representative from each gauge orbit is present. Since the gauge orbits are in a many-to-one representative relationship with distinct instantaneous states instituting a one-to-many relationship between them and points in our space will fix these points as providing a representation with a one-to-one correspondence to our ontology. We thus see that gauge fixing can achieve an almost identical job to reduction. This should be no surprise however since, as discussed in §3.3.3, when made precise any viable gauge fixing methodology will be closely conceptually and mathematically related to symplectic reduction. Thus, gauge fixing does not properly considered constitute a distinct alternative but to reductionism; this point will become significant within the quantum context, to which we now turn.

Chapter 3

Quantisation of gauge Theories

3.1 Geometric quantization

The objective of the geometric quantisation programme (Echeverria-Enriquez and Munoz-Lecanda (1999)) is to find a correspondence between the sets of pairs constituted by: symplectic manifolds $(\mathcal{M}, \Omega)$ together with smooth real functions $C^\infty(\mathcal{M})$, on the one hand; and complex Hilbert spaces $\mathcal{H}$ together with self-adjoint operators $\mathcal{A}(\mathcal{H})$, on the other. We define the full quantisation of a classical system $(\mathcal{M}, \Omega)$ as a pair $(\mathcal{H}_Q, A)$ under certain conditions on $\mathcal{H}_Q$ and the map, A, which takes us between classical and quantum observables. Explicitly we require that: 1) $\mathcal{H}_Q$ is a separable complex Hilbert space. The elements $|\ \psi\rangle \in \mathcal{H}_Q$ are the quantum wavefunctions and the elements $|\ \psi\rangle_{\mathbb{C}} \in \mathbf{P}\mathcal{H}_Q$ are the quantum states where $\mathbf{P}\mathcal{H}_Q$ is the projective Hilbert space; 2) A is a one to one map taking the classical obervables $f \in \Omega^0(\mathcal{M})$ to the self adjoint operators A_f on $\mathcal{H}_Q$ such that: i) $A_{f+g} = A_f + A_g$ ii) $A_{\lambda f} = \lambda A_f \ \forall \lambda \in \mathbb{C}$ iii) $A_1 = Id_{\mathcal{H}_Q}$; 3) $[A_f, A_g] = i\hbar A_{\{f,g\}}$ (i.e., A is a Lie algebra morphism up to a factor); 4) For a complete set of classical observables $\{f_j\}$, $\mathcal{H}_Q$ is irreducible under the action of the set $\{A_{f_j}\}$.

We can see this quantisation programme as consisting essentially of the construction of a Hilbert space $\mathcal{H}_Q$ on which the Lie algebra of classical observables can be represented irreducibly in terms of a set of self-adjoint operators $\mathcal{A}(\mathcal{H}_Q)$ – the elements of this set are the quantum observables. When combined with the symplectic reduction procedure outlined above, geometric quantization gives us a methodology for quantising a system with first class constraints – i.e., first reduce then geometrically quantise the reduced phase space making use of the symplectic structure that our reduction procedure guarantees. Explicitly, what we do is consider the reduced phase space with geometry (Π_R, Ω_R) and set of reduced observable functions $\mathcal{O}_R$ to be our classical pairing and find the corresponding Hilbert space $\mathcal{H}_\mathcal{R}$ and self-adjoint operators $\hat{\mathcal{O}}_\mathcal{R}(\mathcal{H}_\mathcal{R})$. If the symplectic reduction procedure runs through successfully we are 'virtually guaranteed' to be able to construct the quantum equivalent.[17]

3.2 Constraint quantisation

The first step in Dirac's constraint quantisation approach (Dirac (1964); Henneaux and Teitelboim (1992)) is to quantise the extended phase space Γ. As we have seen this space will have a symplectic structure. We can therefore promote smooth phase space functions, $f \in C^{\infty}(\Gamma)$, into Hermitian operators, $\hat{f}$, and the Poisson bracket relation, $\{f, g\} = \Omega(X_f, X_g)$, into commutation relations with the appropriate $i\hbar$ factors. This essentially amounts to a partial application of the geometric procedure above. The Hilbert space that results is called the auxiliary Hilbert space $\mathcal{H}_{aux}$ and we can define a class of auxiliary state vectors $\mid \psi_{aux}\rangle$. We then impose the (first class) constraint functions as operators on $\mathcal{H}_{aux}$ restricting the physical state vectors $\hat{\phi}_j \mid \psi_{phys}\rangle = 0$. The Hilbert space that is constructed by taking the physical states is the physical Hilbert space $\mathcal{H}_{phys}$ of the quantum theory. We are provided with a set of quantum observables $\hat{\mathcal{O}}$ by considering the set of self-adjoint operators which commute with the constraints and map physical states to physical states.

Formally the quantisation procedure we have just sketched suffers from a number of difficulties and ambiguities not least: 1) the quantisation of the classical constraint functions on phase space is not unique due to a factor ordering ambiguity; 2) extra input is needed to define a Hilbert space structure on the physical states in particular an inner product; 3) solving the constraints at the quantum level is non-trivial and may lead to inconsistent results.[18]

[17] There are here numerous qualifications and extra subtleties regarding geometric quantization that might have been discussed in more detail. We have not, for example, ventured into discussion of Van Hove theorem, pre-quantisation polarisation, or other formal features that imply that symplectic structure is not on its own sufficient to guarantee a viable quantization. Although significant in of themselves, such complications are not directly relevant to our investigation and thus their neglection is appropriate. See Gotay (1980); Woodhouse (1997).

[18] There is also the additional problem that if the constraints depend non-polynomially on the field variables then it may prove impossible to find a rigorously defined representation of them on the $\mathcal{H}_{aux}$. This issue is particularly pressing for the constraints of canonical general relativity and leads, in that case, to the introduction of Ashtekar variables. However, neither this formal issue, nor the structure of the new variables, have any particular bearing on the our more conceptual concerns regarding the nature of quantum Hamiltonian constraints. Their discussion can, therefore, be reasonably neglected for the purposes of this non-explicit treatment.

A number of modern strategies are available that allow us to formalise the Dirac quantisation scheme such that these issues can be overcome or at least diminished. The two that we will briefly consider here are the *group averaging methodology* (as used in refined algebraic quantisation) and the *direct integral approach* (as applied in the master constraint programme). These two techniques are particularly significant for our purposes since a combination of them is utilised in the loop quantum gravity approach to quantising general relativity (Thiemann (2007)).

3.2.1 Refined algebraic quantisation

Refined algebraic quantisation (Giulini and Marolf (1999b,a)) (RAQ) is a methodology for addressing the ambiguities of the Dirac quantisation scheme whilst still staying within the broad outline of 'quantise first, constrain second'. As per the original Dirac approach, we first construct a Hilbert space representation of the operator algebra of functions on the extended phase space. The constraints are then taken to be represented as Hermitian operators acting on this $\mathcal{H}_{aux}$. Crucially, we require that the commutator algebra of the quantum constraints forms a Lie algebra (this will always be the case provided the classical Poisson bracket constraint algebra closes with structure constants) – exponentiation of the constraint operators will then yield a unitary representation $U(g)$ of the corresponding Lie group G. Let us then define some subspace $\Phi \subset \mathcal{H}_{aux}$ together with its algebraic dual $\Phi^\star$ (i.e., the space of complex valued linear functions f on Φ). If the space Φ is chosen such that the constraint operators map it into itself then a well defined dual action of these operators is also available.[19] Solutions of the constraints are then elements $f \in \Phi^\star$ for which $U(g)f = f \; \forall g \in G$. Physical observables can then be defined as self adjoint operators $\hat{\mathcal{O}}(\mathcal{H}_{aux})$ which include Φ in their domain, map Φ to itself and (crucially) commute with the group action on Φ.[20]

The pivotal move is the definition of the *rigging map* which is an anti-linear map η from Φ into $\Phi^\star$ such that: its image solves the constraints; it is real and positive; it commutes with the observables. The RAQ scheme then provides us with a methodology for constructing the physical Hilbert space since an inner product is provided to us by the rigging map: $\langle \eta(\varphi_1), \eta(\varphi_2) \rangle_{phys} = \eta(\varphi_1)[\varphi_2]$. This new inner product is defined on Φ and it leads us to the physical Hilbert space $\mathcal{H}_{phys}$ via taking the quotient of Φ by the sub-set of

[19] i.e., we have that $U(g)f[\varphi] = f(Ug^{-1}\varphi), \forall \varphi \in \Phi$

[20] $\hat{\mathcal{O}}U(g) \mid \varphi\rangle = U(g)\hat{\mathcal{O}} \mid \varphi\rangle, \forall g \in G, \varphi \in \Phi$

zero norm vectors it defines. The physical observables will then be automatically defined as operators on $\mathcal{H}_{phys}$ and the correspondence between the RAQ definition of observables and the original Dirac one given above, becomes explicit.

Clearly, the success of RAQ depends on our ability to find a suitably unique rigging map. This can be done subject to the restriction that G is a locally compact Lie group with a Haar measure μ_H.[21] In these circumstances (and for the case that the group is unimodular, see Giulini and Marolf (1999a) for the non-unimodular case) then the group averaging methodology defines the rigging map simply as:

$$\eta \mid \varphi\rangle := \langle\varphi \mid \int d\mu_H U(g) \tag{3.1}$$

That this rigging map solves the constraints is guaranteed by the invariance of the Haar measure and that it is real and commutes with the observables is guaranteed by the fact that it is invariant under $g \to g^{-1}$.

3.2.2 The Master Constraint Programme

The Master Constraint Programme (Thiemann (2006, 2007); Dittrich and Thiemann (2006)) (MCP) for the quantisation of constrained systems constitutes more of a departure from the Dirac scheme than RAQ since it leads us to a different representation of the constraint functions even at a classical level. It is still of the Dirac quantisation genus, however, since these reformulated constraints are again only imposed after quantisation. A particular strength of the MCP approach is that it remains well defined even for systems where the Poisson bracket algebra of the constraints closes only with structure functions. This is particularly important feature for our purposes since the Hamiltonian constraints of canonical general relativity are associated with an algebra of exactly this type.

The essential idea is to re-write the classical constraint functions, $\phi_j(p, q) = 0$, in terms of a single equation which will be satisfied under the same conditions. This new single constraint is then the *Master Constraint*. A simple example is given by taking a positive quadratic two-form K^{ij} and constructing the equation:

$$\mathbf{M} := K^{ij}\phi_i\phi_j = 0 \tag{3.2}$$

[21] A right (left) Haar measure is a positive measure on a group invariant under right (left) translations. For the uni-modular case which we are restricting ourselves to, the left and right Haar measures agree.

This equation is satisfied if and only if all the individual constraint functions are vanishing and thus defines the same physical phase space Σ that we had before. We can recover our observable condition for the *extended phase space* by considering the class of functions such that:

$$\{\{\mathbf{M}, \mathcal{O}\}, \mathcal{O}\}|_{\mathbf{M}=0} = 0 \tag{3.3}$$

i.e., those functions which have a vanishing double Poisson bracket with the master constraint on the constraint surface. The geometric interpretation of this condition on classical observables is subtly, yet importantly, different to the standard one given above. Strictly, it is a restriction that implies that the observable functions generate finite symplectomorphisms which preserve Σ, rather than the usual condition that the observables are constant along the null directions generated by the constraints on Σ. However, it can be straightforwardly demonstrated that the two conditions are equivalent Thiemann (2006). Thus the intuitive connection between these observables and the $\mathcal{O}_R$ of §2.1 is retained. We may therefore think about the $\mathcal{O}$ as corresponding to functions projected up from the reduced phase space.

Moving on to quantisation, we look for a representation of the Poisson algebra of functions on the extended phase space, f, in terms of commutator algebra of operators, $\hat{f}$, on a (separable) kinematic Hilbert space $\mathcal{K}$. We then require that the Master Constraint $\mathbf{M}$ is represented as a positive, self-adjoint operator $\hat{\mathbf{M}}$. This is possible even if the classical constraints cannot themselves be represented in such a way – i.e., when they fail to form a Lie algebra under the Poisson bracket operation. Following Thiemann (2006, 2007), since $\mathcal{H}_{aux}$ is by assumption a separable Hilbert space it can be represented as a direct integral of separable Hilbert spaces $\mathcal{H}^{\oplus}_{aux}(\lambda)$, $\lambda \in \mathbb{R}$, subordinate to $\hat{\mathbf{M}}$ according to:

$$\mathcal{H}_{aux} = \int_{\mathbb{R}}^{\oplus} d\nu(\lambda)\mathcal{H}^{\oplus}_{aux}(\lambda) \tag{3.4}$$

where although the measure ν and Hilbert spaces $\mathcal{H}^{\oplus}_{aux}(\lambda)$ are not uniquely determined, different choices will give rise to unitarily equivalent Hilbert spaces. Crucially, we can show that, for such a direct integral decomposition, we will have that $\hat{\mathbf{M}}$ acts on $\mathcal{H}^{\oplus}_{aux}(\lambda)$ by multiplication by λ. We can then define a physical Hilbert space $\mathcal{H}_{phys} := \mathcal{H}^{\oplus}_{aux}(0)$ which automatically comes equipped with a well defined inner product and upon which we can consider, if the uniform limit exists, a prospective class of observables in terms of the *ergodic mean* of the $\hat{f}$, :

$$[\hat{f}] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} dt e^{it\hat{\mathbf{M}}} \hat{f} e^{-it\hat{\mathbf{M}}} \tag{3.5}$$

The existence of this object is guaranteed by the Birkhoff ergodic theorem since the unitary evolution operator, $U(t) = e^{it\hat{\mathbf{M}}}$, is a one-parameter measure preserving transformation on the Hilbert space (see Walters (1981, §1.6)). From this definition we have that the $[\hat{f}]$ will both preserve the physical Hilbert space and induce a self-adjoint operator on that space. Furthermore, provided the spectral projections of the bounded operator $[\hat{f}]$ commute with those of $\hat{\mathbf{M}}$ (which we may expect) the ergodic mean can be seen to constitute a member of the class of *strong observables*, $\hat{\mathcal{O}}_s$. These observables are defined to be functions such that relevant commutator with the master constraint vanishes identically: $[\hat{\mathcal{O}}_s, \hat{\mathbf{M}}] \equiv 0$. Strong observables form a sub-set of the *weak* observables $\hat{\mathcal{O}}$ defined via the quantum equivalent of the condition given by equation (3.3).

It is important to note that despite the impressive improvements in formalising the Dirac quantisation programme that the RAQ and MCP approaches enable, these advances have come at the cost of removing our quantum formalism far from physical intuition. The most obvious way to ensure we have constructed more than just a mathematical edifice would be to demonstrate that both schemes have the appropriate classical limit. In particular, a proof that the quantum observables reproduce their classical analogues in the appropriate limit would be highly desirable. Alternatively, one might seek to anchor these 'quantise first, reduce second' techniques by a formal, or at least conceptual, correspondence with the less intuitively opaque 'reduce first, quantise second' alternatives. It is to this task will turn in §4.1.

3.3 Path integral quantisation

3.3.1 Feynman path integral quantisation in phase space

Before we introduce the Faddeev-Popov methodology for the path integral quantisation of a standard gauge theory, a few brief remarks concerning the origins of the Feynman path integral formulation of quantum theory are necessary. We will confine ourselves here to particle mechanics. Expressions corresponding to the extension of these ideas to infinite dimensional field theories can be found in (for example) Peskin and Schroeder (1995).

Consider an unconstrained classical system with $S[q(t)] = \int_\gamma \mathcal{L}(q, \dot{q})dt$ defined in the usual $6n$-dimensional velocity-configuration space. We can take the path γ as defining an arbitrary path in *configuration space* between an initial configuration q_i and a final configuration q_f. Classically we know that only one such path is physically realisable - that corresponding to extremisation of the action. However, in the quantum realm we know that there can be multiple paths that can be physically realised - this is manifestly demonstrated by real world quantum systems such as that described by the famous two split experiment. There, not only are multiple paths found to be possible given a fixed identical configuration but also for fixed initial *and final* configurations - we can block either slit and still get a detector reading in the same place. In this case, and in general we know - by the superposition principle - that the total amplitude for any process is equal to the coherent sum of the amplitudes for each of the possible ways this process can be realised. It is precisely for the calculation of this total amplitude that the path integral expression is designed.

From standard Schrödinger quantum mechanics in the position representation we have that the time evolution operator is simply $e^{\frac{-iHT}{\hbar}}$ where T is some finite time. Thus we can write the amplitude for transition between a initial positional state $\mid q_i\rangle$ and a final positional state $\mid q_f\rangle$ as $\langle q_f|e^{-\frac{i}{\hbar}H(t_f-t_i)}|q_i\rangle$. If we write the amplitude for each path as a pure phase the superposition principle leads us to consider the intuitively sensible heuristic expression:

$$\langle q_f|e^{-\frac{i}{\hbar}H(t_f-t_i)}|q_i\rangle = \sum_{\text{all paths}} e^{i.(\text{phase})} \tag{3.6}$$

Getting from here to the full Feynman path integral expression involves two steps corresponding to replacing 'phase' and 'sum over all paths' with mathematics such that the result coheres with experiment. We achieve the first via the physically well motivated postulation that since $e^{\frac{i}{\hbar}S[q(t)]}$ will give us the appropriate classical limit it is the correct expression for the quantum amplitude of a path.[22] The second requires us to introduce the formal machinery of functional integration. Considering the path integral between two points q_f and q_i we break the time interval $T = t_f - t_i$ into discrete infinitesimal pieces ϵ and label the spatial coordinates of each successive slice by the suffix K which runs up

[22]This is Feynman's second postulate in his original derivation of the path integral formulation (Feynman (1948)) and derives its origin from remarks due to Dirac (1933) concerning the relation between the classical action and quantum theory. The framing of this postulate in terms of a rigorous conceptual basis is an outstanding and intriguing question which unfortunately falls outside the remit of our current project.

to $N-1$. We can then define the sum over all paths in an intuitively sensible manner as:

$$\sum_{\text{all paths}} = \int \mathcal{D}q(t) \equiv \frac{1}{C(\epsilon)} \int \frac{dx_1}{C(\epsilon)} ... \int \frac{dx_{N-1}}{C(\epsilon)} = \frac{1}{C(\epsilon)} \prod_{K}^{N-1} \int_{-\infty}^{\infty} \frac{dx_k}{C(\epsilon)} \tag{3.7}$$

Following Peskin and Schroeder (1995) we can fix the constant $C(\epsilon)$ for the case of a particle mass m moving in a one dimensional potential by considering the particular infinitesimal time slice in which t goes from $T-\epsilon$ to T. Sending ϵ to zero, expanding in a power series and then performing the necessary Gaussian integrals allows us to show that:

$$C(\epsilon) = \sqrt{\frac{2\pi\hbar\epsilon}{-im}} \tag{3.8}$$

and therefore obtain an explicit form of the path integral (for the one particle case):

$$\langle q_f | e^{-\frac{i}{\hbar}H(t_f - t_i)} | q_i \rangle = \lim_{\epsilon \to 0} \left(\frac{m}{2i\pi\hbar\epsilon} \right)^{\frac{N}{2}} \prod_{K}^{N-1} \int_{-\infty}^{\infty} dx_k e^{\frac{i}{\hbar} \int_0^T \mathcal{L}(q,\dot{q}) dt} \tag{3.9}$$

which is such that it can be explicitly shown to reproduce normal Schrödinger evolution. We can express this relationship between the path integral formulation and the Schrödinger formulation of quantum mechanics by writing:

$$U(q_i, q_t; T) = \langle q_f | e^{-\frac{i}{\hbar}HT} | q_i \rangle = \int \mathcal{D}q(t) e^{\frac{i}{\hbar} S[q(t)]} \tag{3.10}$$

$$i\hbar \frac{\partial}{\partial T} U(q_i, q_t; T) = U(q_i, q_t; T) \tag{3.11}$$

This path integral expression describes quantum mechanical behaviour in a configuration space. For our purposes we need a more general expression corresponding to a path integral for phase space. This requires us to consider a *functional measure* relevant to this space rather that that for configuration space which we have just considered. Above we took the configuration space functional measure $\mathcal{D}q(t)$ to be such that it weighted the contribution from each infinitesimal section of each path equally. We take a similar approach to arrive at a functional measure for phase space and make the hypothesis that we should be looking for a $\mathcal{D}q(t)\mathcal{D}p(t)$ such that the integration will extend over all physically distinct configurations, and weight each by the same trivial factor of unity

(Unz (1986)). This means that the phase space path integral,

$$\langle q_f|e^{-\frac{i}{\hbar}H(t_f-t_i)}|q_i\rangle = \left(\prod_i \int \mathcal{D}q(t)\mathcal{D}p(t)\right) e^{\frac{i}{\hbar}\int_{t_i}^{t_f} dt(\sum p_i\dot{q}^i - H(q,p))} \tag{3.12}$$

will have the general and simple form:

$$\langle q_f|e^{-\frac{i}{\hbar}H(t_f-t_i)}|q_i\rangle = \prod_{i,t}\int \frac{dp(t)_i dq(t)_i}{(2\pi\hbar)^n} e^{\frac{i}{\hbar}\int_{t_i}^{t_f} dt(\sum p_i\dot{q}^i - H(q,p))} \tag{3.13}$$

Since our view as to its meaning is essential in the discussion that follows let us briefly consider the interpretational consequences of using this expression as the fundamental description of a physical system. In a classical system we take phase space to provide us with a representation such that points correspond to instantaneous physical states and curves correspond to dynamical histories. Dynamical histories correspond to curves such that they are the integral curves of the Hamiltonian vector field and we can thus think of the Hamiltonian as generating physical evolution between instantaneous states. Quantum mechanically instantaneous physical states are represented by vectors in a Hilbert space and dynamical evolution is represented in terms of a unitary operator in that space and, as assumed above, in general this operator takes the form $e^{\frac{-iHt}{\hbar}}$ where H is the quantum mechanical Hamiltonian operator and t is the time parameter.

What the canonical path integral expression gives us is the probability for transition between two quantum mechanical states in the position basis, $|q_i\rangle$ and $|q_f\rangle$ – there is no inherent temporal ordering in this transition but we can label the states *initial* and *final* for practical convenience. The construction of the path integral is such that this transition probability is calculated by considering possible transition through each distinct unit of *classical* phase space weighted by the exponential of its quantum action. Thus, in effect what we are doing is considering *every* classical phase space curve – i.e., those corresponding to dynamical histories and those not – and then applying a quantum weight such that the probabilities will match those of normal Schrödinger evolution. Thus there is a sense in which the representation of dynamics in this formalism rests upon a form of ontological equality between each and every instantaneous physical state as represented classically by points in phase space. In of itself this is an intriguing situation that warrants deep and careful philosophical analysis. However, such an investigation does no accord with the task at hand which is a philosophical analysis of the interpretation and quantisation of gauge theory. Rather what concerns us is the degree to which this implied ontological equality is transferred over to systems whose phase space representation is of the constrained and degenerate form discussed in the previous chapters. To frame this question we will introduce a path integral formulation designed for canonical gauge theories – that due to Faddeev and Popov (1967).

3.3.2 Faddeev-Popov path integral quantisation

Let us consider a generic constrained Hamiltonian theory with extended phase space a symplectic manifold (Γ, Ω) and physical phase space a presymplectic manifold (Σ, ω) defined by satisfaction of the set of set of first class constraints, $\phi_\alpha(q,p) = 0$ for $\alpha = 1, ..., m$. A naive path integral quantisation of this theory simply using the general canonical expression defined in the last section applied to Γ would rest upon us assuming a equality between classical phase space points. Since the set of points in the complement of Σ relative to Γ are unphysical this naive approach approach is clearly incorrect. Furthermore, even a path integral purely defined upon Σ will still not give us a sensible result since it would involve counting as representing distinct physical states points which lie on the same gauge orbit and are therefore only different up to the unphysical gauge transformations. Rather we need to adjust the measure in our path integral such that we are only counting with respect to a unique representation of the *classical* ontology.

One consistent way of doing this, as was realised by Faddeev and Popov, is to change the functional measure such that rather than integrating over all possible phase space points we instead integrate over a subspace, Π_{GF}, of the physical phase space which is such that it intersects each gauge orbit exactly once. Such an integral will only count one point out of each gauge orbit as representing a distinct ontological object and would therefore be expected to lead us to a quantum theory in which only the physical classical degrees of freedom have been quantised. Thus in intent this approach is closely related to the reductive quantisation methodologies discussed previously. We will explore this interpretive connection more fully once we have introduced the necessary mathematical formalism.

First of all, we need to define our subspace. This can be done by first considering a *gauge fixing sub-manifold*, Δ, within the extended phase space by satisfaction of the conditions $\chi_\alpha(q,p) = 0$ with i) $\{\chi_\alpha, \chi_\beta\} = 0$ and ii) $det||\{\chi_\alpha, \phi^\beta\}|| \neq 0$. The condition i) will prove crucial to our ability to introduce convenient canonical co-ordinates and the condition ii) fixes the geometrical relationship between Δ and the gauge orbits on Σ in a precise manner - which we will detail shortly.

The intersection of Δ and Σ is then exactly the subspace, Π_{GF}, we are looking for. Since we know that m first class constraints equate to $2m$ excess degrees of freedom in the extended phase space, then since the dimension of Π_{GF} is $(2n - m) - m = 2n - 2m$, it clearly has the correct dimensionally. Our second condition on the $\chi_\alpha(q, p) = 0$ can then be expressed in terms of Π_{GF} being a transversal integral manifold of the distribution of zero norm vectors defined by the kernel of the presymplectic form ω on Σ (Faddeev 1969). This imposes a condition on χ_α and ϕ^β such that the gauge fixing only selects a single member of each gauge orbit and thus that Π_{GF} is nowhere parallel to the null vectors which define the gauge direction. Further to this we also impose the global requirement; iii) Π_{GF} intersects the gauge orbits exactly once. Together ii) and iii) ensure that the space's representational relationship with instantaneous states is uniquely defined - exactly one point per gauge orbit is present so exactly one physically distinct state is represented.

The kernel of brilliance behind the Faddeev-Popov methodology (FPM) (Faddeev and Popov (1967); Popov and Faddeev (2005); Popov (2010)) is to define the relevant path integral in Γ leaving the specification of Δ – beyond our conditions i), ii) and iii) – free. This allows us to implicitly consider a standard canonical path integral in Π_{GF} in a reduced set of canonical coordinates without having to explicitly construct this space by fixing the χ_α. The important step is the definition of the functional measure $d\mu(q(t), p(t))$ in the expression:

$$\langle q_f | e^{-\frac{i}{h} H(t_f - t_i)} | q_i \rangle = \int e^{\frac{i}{h} \int_{-\infty}^{\infty} dt (\sum p_i \dot{q}^i - H(q,p))} \prod_t d\mu(q(t), p(t)) \tag{3.14}$$

The Faddeev-Popv Ansatz is that it should take the form:

$$d\mu(q(t), p(t)) = \prod_{i,\alpha} \delta(\chi_\alpha) \delta(\phi^\alpha) det || \{\chi_\alpha, \phi^\beta\} || \frac{dp_i(t) dq_i(t)}{(2\pi h)^{n-m}} \tag{3.15}$$

Each element of this can be justified on an intuitive basis. The two functional delta functions restrict our integration to the gauge fixing surface and constraint surface respectively – we can think of each of them as an infinite product of delta functions, one for each phase space point. The determinate then gives a weighting factor (which we know by ii) to be non-zero) based upon the geometrical relationship between Δ and the gauge orbits associated with the constraints. The precise geometrical basis behind this *Faddeev-Popov determinate* is complex but can in fact be traced to the metric structure of the reduced phase space discussed earlier. For details see Babelon and Viallet (1979); Ordóñez and Pons (1992).

Combing this functional measure with our general path integral expression then gives:

$$\int e^{\frac{i}{h}\int_{-\infty}^{\infty} dt(\sum p_i \dot{q}^i - H(q,p))} \prod_{i,t,\alpha} \delta(\chi_\alpha)\delta(\phi^\alpha) det||\{\chi_\alpha, \phi^\beta\}|| \frac{dp_i(t)dq_i(t)}{(2\pi h)^{n-m}} \tag{3.16}$$

which is the Faddeev-Popov path integral for a constrained Hamiltonian theory subject to gauge fixing conditions χ_α.

3.3.3 Classical reduction and Faddeev-Popov quantisation

As has been suggested by a number of remarks above the FPM can understood as being grounded upon a distinct stance as to the representational structure of classical gauge theory. In particular, as highlighted above, precisely what we are looking for in the restrictions we make on the gauge fixing functions and our construction of the measure is a methodology for picking out a sub-manifold within which exactly one representative from each gauge orbit is present. Since the gauge orbits are in a many-to-one representative relationship with distinct instantaneous states instituting a one-to-many relationship between them and points in our space Π_{GF} will fix these points as providing a representation with a one-to-one correspondence to our ontology. Thus, the basis of the F-P path integral could be argued to rest upon a reductive interpretation applied to the classical theory – in that Π_{GF} has an identical representational role to the reduced phase space, Π_R, which we have defined previously. An immediate question is then whether we should confer upon the FPM some degree of formal or representational equivalence with the *reduce first, quantise second*, geometric approach discussed above.

Thanks to some elegant work by Faddeev (1969), one can in fact explore the relationship quite clearly by introducing a new set of canonical coordinates on phase space which are such that Π_{GF} is itself canonical coordinatised.[23] Continuing with the notation above, let us define the new set of canonical coordinates on Γ as $\{q^\alpha, Q^a, p_\alpha, P_a\}$. We require

[23] That such a set of coordinates is always available is guaranteed by the condition i) imposed on the χ_α above.

that they are are such that $\chi_\alpha(q,p) = p_\alpha$. Clearly $a = 1, ..., r$ where $r = n - m$. Since we have that $det||\frac{\partial\phi^\alpha}{\partial q^\beta}|| \neq 0$ we can solve the constraints for the q^α and define Π_{GF} in terms of the expressions $p_\alpha = 0$ and $q^\alpha = q^\alpha(Q^a, P_a)$. We will then have that (Q, P) are independent coordinates on Π_{GF}.

We can now demonstrate a pair of fundamental and powerful equivalences. Firstly, consider a standard canonical Feynman path integral of the form

$$\int e^{\frac{i}{h}\int_{-\infty}^{\infty} dt(\sum P_a \dot{Q}^a - H(Q,P))} \prod_{a,t} \frac{dP_a(t)dQ_a(t)}{(2\pi h)^r} \tag{3.17}$$

This expression gives us a quantum theory based upon the classical physical phase space Π_{GF} and is thus equivalent to a geometric quantisation of that space – at least to the extent that path integral quantisation is equivalent to geometric quantisation in general (see relevant discussion in §16.3). Furthermore, as shown by Faddeev (1969), this expression is formally equivalent to one of the standard Faddeev-Popov form – i.e.,

$$\int e^{\frac{i}{h}\int_{-\infty}^{\infty} dt(\sum p_i \dot{q}^i - H(q,p))} \prod_{t} d\mu(q(t), p(t)) \tag{3.18}$$

with the measure as defined above. This means that application of the FPM is equivalent to geometric quantisation of the space Π_{GF} (again up to the equivalence between path integral and geometric methods).

Next we can show explicitly that functions on Π_{GF} are connected to those on Π_R. As discussed above, we can associate every $f_R \in C^\infty(\Pi_R)$ with an observable phase space function $\mathcal{O}$ which is such that $\{f, \phi_\alpha\}|_\Sigma = 0$. Now crucially, according to Faddeev (1969 pp.4-5), a Poisson bracket of phase space functions evaluated on Σ is equivalent to that defined with respect to the new canonical functions on Π_{GF}. Expressing these functions explicitly as $f_{GF} = f(q^\alpha(Q^a, P_a), Q_a, 0, P_a)$ this means that we have that:

$$\{f, g\}|_\Sigma = \sum\left(\frac{\partial f_{\Pi_{GF}}}{\partial P_a}\frac{\partial g_{\Pi_{GF}}}{\partial Q_a} - \frac{\partial f_{\Pi_{GF}}}{\partial Q_a}\frac{\partial g_{\Pi_{GF}}}{\partial P_a}\right) \tag{3.19}$$

where $f_{\Pi_{GF}} = f(q^\alpha(Q^a, P_a), Q_a, 0, P_a)$. We can therefore assert that there is a fundamental connection between the Poisson bracket algebra of classical observables defined within the reduced and gauge fixed formalisms. In fact, we can make the specific statement that between the algebra's there exists a *symplectic isomorphism*. Such connections will become important to our discussion in Part IV and we will refer back to the philosophical importance of this result then.

Chapter 4

Quantisation, reduction and ontology

4.1 Representative commutation between quantisation and reduction

Provided we assume that the Poisson bracket algebra of the classical constraints closes with structure constants (and several further restrictions are imposed), we can often prove formally that a Dirac type procedure of quantising and then imposing the constraints at the quantum level is equivalent to first symplectically reducing and then geometrically quantising. This is a specific case of what is commonly referred to as the Guillemin-Sternberg conjecture Guillemin and Sternberg (1982). For examples of proofs of the conjecture for various degrees of generality see Gotay (1986); Duistermaat *et al.* (1995); Conrady and Freidel (2009). The crucial results established in such *commutation* proofs is that: i) the physical Hilbert space constructed through a Dirac type approach, $\mathcal{H}_{phys}$, can be shown to be unitarily isomorphic to that (i.e., $\mathcal{H}_R$) achieved by quantising the symplectic manifold constructed by a classical reduction of the action of the constraints; and ii) the two quantization procedures result in an equivalent set of observables to the extent that the isomorphism in i) also entwines the representations of the two sets of quantum observables (both of which can be connected back to the same set of gauge invariant classical observables). We can thus assert, in certain circumstances, that quantisation does commute with reduction and assert physical equivalence in a strict sense.

What will be important for our discussion below is that the constraints of canonical general relativity are of such formidable complexity that the theory lies well outside any of the existent commutation proofs. With this future issue in mind we will introduce here a conceptual notion of commutativity which can be established in cases where formal arguments are not available. Such a weaker notion of commutativity will allow us to tackle the important task of exploring the somewhat unclear conceptual foundations of quantum theories constructed via a Dirac type methodology. In particular, it will give us a basis in which to examine the extent to which mathematical structures which they present us with are in correspondence with a reasonable physical ontology.

To establish a conceptual notion of commutativity we do not need to look for the existence of a suitable observable entwining isomorphism existing between $\mathcal{H}_{phys}$ and $\mathcal{H}_R$ but rather establish correspondence between both the Hilbert spaces and the observables on a representative level. We can flesh this idea out in terms of how the two approaches treat the relevant symmetries, observables and degrees of freedom of a given theory. What interests us is the extent to which the imposition of the constraints at a quantum level should be understood as implementing the same reduction from an otiose to a unique representative structure that we enact via classical reduction. The key features of the classical symplectic reduction procedure that we must require to be replicated at a quantum level are: 1) quotienting by the same gauge group; 2) reduction by the same number of degrees of freedom; and 3) the quantum observables defined via the two routes are equivalent to the extent that we are justified in thinking of them as representing the same underlying ontology. If we are satisfied as to equivalence in these three senses then we are justified in asserting that the two quantisation procedures produce *representationally equivalent* structures and that representative commutation between quantisation and reduction holds. Clearly, this notion of commutativity in terms of representative equivalence should be implied by that defined in terms of unitary isomorphism but not visa versa. However, its potential significance is not *purely interpretive* since, given a case where we believe the classical reduction to lead to a physically unrealistic reduced phase space, establishing representative commutation will then provide us with grounds to doubt the physical basis of the theory quantised along the Dirac lines. Thus, what we have constructed is a heuristic tool as well as an interpretative criterion of equivalence.

Let us consider the case of a theory where: i) the constraints are associated with a Poisson bracket algebra with structure constants such that we can represent their action quantum mechanically via a set of unitary operators on an auxiliary Hilbert space; and ii) classically we can construct a reduced phase space with a symplectic geometry and a non-trivial Hamiltonian operator. In these circumstances, we can apply RAQ to produce a quantum theory via the Dirac type approach or alternatively proceed with a geometric quantisation of the reduced space. The key to evaluating our notion of representative commutativity is to examine the degree of correspondence between the quantum and classical reduction procedures. Immediately, we can see a *prima facie* correspondence between the classical gauge orbits defined by the constraints on the physical phase space and the *quantum orbits* defined by the action of the unitary representation of the Lie group generated by the constraints, $U(g)$, on the auxiliary Hilbert space. Further to this, we can also see an intuitive correspondence between a) the quotienting of the orbits in the classical theory to enable passage to the reduced space and b) the group averaging over the *quantum orbits* that is used to construct the rigging map which projects into the physical Hilbert space. However as pointed out by Corichi (2008) we must be wary of taking these resemblances too seriously. Unlike in the classical case, the orbits are not generically equivalence classes of physical states – this is to be expected since in the quantum case we do not make any restriction to a physical yet degenerate sub-space of the auxiliary Hilbert space which would be analogous to the physical phase space. Furthermore, the rigging map defined by group averaging will – unlike the map to the reduced phase space – in general take us to a state which is not part of the *quantum orbit*. Thus, the two quotienting procedures are clearly different in an important sense.

Nevertheless, despite these differences the two procedures *are* equivalent in terms of quotienting out the same gauge group. Since the constraints form a Lie algebra they are associated with a Lie group, G, at a classical level. It is the action of this group that we are removing from the physical phase space via symplectic reduction. This is the same group that we represent in RAQ in terms of unitary operators on $\mathcal{H}_{aux}$ and that we quotient out via the rigging map defined by group averaging in order to construct $\mathcal{H}_{phys}$.

To see the correspondence in terms of degrees of freedom reduction we have to consider the nature of the group averaging procedure a little more carefully. Clearly, if all we were doing in RAQ was the quantum equivalent of classical reduction on an unconstrained space then we would have a mismatch in terms of number of degrees of freedom removed – in the classical procedure half the excess degrees are removed by restriction to the physical phase space and half by the reduction itself. Rather, we must be able to understand the group averaging procedure as achieving both steps at once. We have, in fact, already considered the essence of the answer – the rigging map does not just reduce out equivalence classes it projects onto physical states. If we start out with an unphysical state then it will take us to a physical state. If we start out with a solution to the constraints then, because the orbit is trivial, group averaging will keep us at the same point. Thus, as well as quotienting out the same gauge group we also have the quantum equivalent of restriction to the physical phase space and the desired correspondence in terms of degrees of freedom reduction is guaranteed.

Since the quantum observables of the RAQ scheme are defined such that they commute with the group action on the relevant sub-space of $\mathcal{H}_{aux}$, there is also a clear intuitive relationship between them and the classical observables on the physical phase space $\mathcal{O}$: both are in a sense 'constant along the gauge orbits' – although of course as we have seen the quantum gauge orbits are of a very different character to the classical ones. A stronger relationship can be established between the observables on the reduced phase space $\mathcal{O}_R$ and the quantum observables of RAQ since both are well defined with respect to a non-degenerate and physical representative structure (i.e., Π_R and $\mathcal{H}_{phys}$). Moreover, since there is a correspondence between the way *states* are represented in Π_R and $\mathcal{H}_{phys}$ respectively, there is also a direct connection between the representational roles of the observables: in each case points in the respective spaces can be given analogous representative roles as corresponding to unique instantaneous states, and this then means we can establish a relationship between the observables defined via functions/operators defined over the state spaces. In this context, we can then consider the associated reduced quantum observables $\hat{\mathcal{O}}_R$ defined on $\mathcal{H}_{\mathcal{R}}$ to precisely parallel (in a representational sense) the RAQ observables when defined simply as operators on $\mathcal{H}_{phys}$.

It would therefore seem clear for the class of theories within which RAQ and symplectic reduction are applicable, representative commutativity of reduction and quantisation will hold. The important question of whether our condition also holds for theories within which RAQ is not applicable, and the master constraint programme for quantisation has been applied, will be considered in the context of the Hamiltonian constraints of canonical general relativity in §14.2. For the moment we will focus our discussion upon the interpretational consequence which we can attach to the establishment representative commutation for standard gauge theories.

4.2 Quantisation and reductionism

In §3.3.3 we detailed how and why the Faddeev-Popov quantisation methodology can be understood in reductive terms – in effect, when we consider a Faddeev-Popov path integral we are considering a Feynman path integral on a reduced phase space. This means that Faddeev-Popov quantisation is equivalent to reduced phase space quantisation up to the (incomplete) equivalence between geometric and Feynman path integral quantisation.

In the previous section we established an argument that in the wide class of theories for which the RAQ refinement of Dirac quantisation is available, such a route towards quantisation should be understood as representationally equivalent to a quantisation of the reduced space. Together one might see these details as giving motivation for us to argue that the quantisation of standard gauge theory in general is predicated upon the quantisation of the classical reduced phase space. To make such an argument would, in effect, to be to argue that the doctrine of classical reductionism is in some way always implicit in the quantisation of standard gauge theories.

There are a number of problems with this argument from the structure of quantisation to reductionism[24]. First, since the classical and quantum theories are strictly distinct, a quantum based argument in favour of the reductionist stance does not necessarily impinge upon the classical debate. Second, there exist quantisation methodologies for standard gauge theories that go beyond those we have considered so our analysis is in this sense incomplete. Furthermore, it is arguable that, given the example of the BRST technique (Henneaux and Teitelboim (1992)), such additional techniques may amount to enlarging rather than reducing the phase space. Third, when non-reductive techniques lead to the same quantum theory as reductive techniques we could also argue that the quantum formalism that they lead to is the fundamental one and so that anti-reductionism is implicit in quantization. Fourth, to the extent that the two methodology types lead to formally different quantum theories there is still scope (over and above representative equivalence) to argue that one may be true and one false – and it may in fact be an empirical matter to decide which one is correct.

With regard to the second of these points, one may argue that BRST technique is in essance a more mathematically sophisticated version of Faddeev-Popov technique – where the F-P determinate is expanded in terms of *ghost fields* – and therefore that interpretationally BRST too can be understood in terms of a Feynman path integral on the reduced space. Thus, detailed consideration may allow the rebuttal of at least one of the criticisms of our argument. However, that still leaves three more, and it seems unlikely that a convincing answer can be found to them all. Rather, we might more reasonably assert that quantisation merely transfers the reductionism/anti-reductionism debate to the another level – there is no fundamental sense in which it can be evoked to settle it (this situation of *underdetermination* in standard gauge theory will be discussed in Chapter 17).

On the other hand, although problematic, the idea of connecting quantisation to reductionism undoubtably latches onto a key kernel of truth. In particular, there seems ample

[24]Many of these points are found in a similar form within Rickles (2008).

scope for running a negative rather than positive argument. For a case in which (classical) reductionism proves to be philosophically or conceptually problematic, we have good cause to re-examine any quantisation technique which implicitly endorses its viability. If reductionism proves incoherent and the reduced phase space fails to correctly parameterise the classical system then any quantisation procedure which is equivalent to quantising the reduced space will rest on dubious foundations. This argument will prove crucial when we are considering the quantisation of a non-standard gauge theory in Chapter 9.

Part II

The Non-Relativistic Problem Of Time

Guide to Part II

In the second part of this thesis we will make our first foray into the class of non-standard gauge theories within which time no longer plays the role of an external background parameter. The particular case which we shall consider is that of non-relativistic reparameterisation invariant theories. Such theories are represented canonically in terms of a constrained Hamiltonian formalism within which the Hamiltonian itself is a first class constraint. As we shall see below, it this feature that when combined with the application of standard gauge theoretic techniques leads to the most acute form of the non-relativistic problem of time: the disappearance of dynamics at both classical and quantum levels. The key to avoiding this acute problem is the insistence that novel techniques must be applied to the theories in question. Detailing the structure and interpretation of such non-standard techniques will be one of the major preoccupations of the following five chapters. As well as its evident intrinsic value for the understanding of the physically interesting models at hand, our non-relativistic investigation shall prove an invaluable foundation for the analysis of the relativistic case in part three.

The basic structure of Part II is as follows. We begin in Chapter 5 by first giving both a Lagrangian (§5.1) and Hamiltonian (§5.2) description of an extended version of mechanics where time is promoted to a configuration variable. We then proceed to the Jacobi formulation of mechanics within which temporal structure is eliminated altogether (§5.3). The final, and crucial, section of Chapter 5 (§5.4) will then introduce an argument against the applicability of standard gauge techniques (in particular symplectic reduction) to the case in hand. Chapter 6 introduces two non-standard strategies for representing time and observables within non-relativistic reparameterisation invariant theory: the emergent time strategy (§6.1) and the correlation strategy (§6.2). Chapter 7 will then consider some of the key interpretational consequences of our discussion in general (§7.1) and of these strategies in particular (§7.2). Chapter 8 will offer some general and introductory ideas about the nature of time in conventional approaches to quantum theory, before we proceed, in Chapter 9, to the introduction of quantum mechanical non-relativistic problem of time. This problem will be seen to have two facets, the first stemming from the reduction issue (§9.1) and the second for the specific exclusion of quantum relational clocks by standard quantisation techniques (§9.2). The nature of these problems is further illustrated by toy model examples (§9.3). We then offer a new proposal for the quantisation of (globally) reparameterisation invariant theories via an intermediary formalism (§9.4) which we will argue to solve the non-relativistic quantum problem of time as we have defined it. We will then consider the structure of the observables of the intermediary formalism in order to demonstrate their ability to be interpreted as representing the *physical* degrees of freedom on the original theory (§9.5), , before finally discussing some outstanding interpretive issues implied by the ideas introduced (§9.6.

Chapter 5

Reparameterisation invariant mechanics

5.1 Extended Lagrangian mechanics

The description of mechanics and gauge symmetry given thus far has made use of a distinguished background parameter; time t. Within the Lagrangian scheme this parameter was associated with both the tangent vectors or velocities, $\dot{q} = \frac{\partial q}{\partial t} \in T\mathcal{C}_0$, and with the preferred parameterisation of the solutions, $\gamma_{PS} : \mathbb{R} \to T\mathcal{C}_0$. An alternative methodology for constructing a mechanical theory is to instead treat time as an additional coordinate, $q_0 = t$, in a $n+1$ dimensional **extended configuration space**, $\mathcal{C} = \mathbb{R} \times \mathcal{C}_0$. Velocities in this space are then defined for all of the $q_\mu \in \mathcal{C}$ by differentiation with respect to an arbitrary parameter τ so we have that $q'_\mu = \frac{dq_\mu}{d\tau}, (q_\mu, q'_\mu) \in T\mathcal{C}$. This arbitrary parameter is also taken to vary monotonically along curves in extended configuration space, $\gamma : \mathbb{R} \to T\mathcal{C}$. Following Lanczos (1970, §5)[25] we can use an extended Lagrangian , $L_{ex}[q_\mu, q'_\mu] : T\mathcal{C} \to \mathbb{R}$ to define an action of the form:

$$I = \int_\gamma d\tau L_{ex}[q_\mu, q'_\mu] = \int_\gamma d\tau (\frac{\bar{T}}{q'_0} - q'_0 V) \quad (5.1)$$

where $\bar{T} = q_0'^2 T$ and all masses are set to unity.

An important property of the extended Lagrangian is that it is homogenous of degree one in the extended set of velocities q'_μ: for some positive number k the transformation $q'_\mu \to kq'_\mu$ implies $L_{ex}[q_\mu, q'_\mu] \to kL_{ex}[q_\mu, q'_\mu]$. This means that the action of our theory will be invariant under re-scalings of the parameter τ. Theories which display such a dynamic insensitivity to parameterisation are said to be **reparameterisation invariant**. The interpretation of this theory will be non-standard since reparameterisation is a symmetry

[25] Also see Johns (2005, §11-12) and Rovelli (2004, §3.1)

of the action which maps between distinct solutions in the extended configuration space – this is because the velocities are parameterisation dependent. Thus these solutions cannot be used to provide a straightforward characterisation of physical histories as in §1.2.

5.2 Extended Hamiltonian mechanics

In correspondence with §1.3 we can define an extended phase space as the cotangent bundle to our extended configuration manifold, $(q_\mu, p_\mu) \in \Gamma = T^*\mathcal{C} = T^*(\mathbb{R} \times \mathcal{C}_0)$, with $p_\mu = \frac{\partial L_{ex}}{\partial q'_\mu}$. The relevant Hamiltonian functional, $H_{ex}[q_\mu, p_\mu] : \Gamma \to \mathbb{R}$ takes the form:

$$H_{ex}[q_\mu, p_\mu] = p^\mu q'_\mu - L_{ex}[q_\mu, q'_\mu] \tag{5.2}$$

which is homogenous of degree one in the set of extended velocities and defines a reparameterisation invariant action

$$I = \int_\gamma d\tau (p^\mu q'_\mu - H_{ex}[q_\mu, p_\mu]) \tag{5.3}$$

By definition we have that the momentum conjugate to time is:

$$p_0 = \frac{\partial L_{ex}}{\partial q'_0} = L_0 - \frac{\partial L_0}{\partial \dot{q}_i} \frac{q'_i}{t'} = -H_0 \tag{5.4}$$

which means the extended Hamiltonian is equivalent to:

$$\begin{aligned} H_{ex}[q_\mu, p_\mu] &= t'(p_0 + H_0) \\ &= 0 \end{aligned} \tag{5.5}$$

The Hamiltonian is therefore a (first class primary) constraint and the dynamics of our theory will be defined upon a surface within extended phase space, $\Sigma = \{x \in \Gamma : H_{ex}(x) = 0\}$. The geometry of the constraint surface is given (as above) by taking the restriction of the relevant Poincaré one form, $\theta = p_\mu dq^\mu$, to Σ:

$$\theta_{|\Sigma} = p_i dq^i - H_0 dt \tag{5.6}$$

and taking the total differential to get a two form $\tilde{\omega} = \mathbf{d}(\theta_{|\Sigma})$ with highly non-trivial structure.[26]

Significantly, this two form is closed and degenerate. Thus the dynamics of extended mechanics is framed within a presymplectic geometry, $(\Sigma, \tilde{\omega})$. That this should be the case can be seen quite simply since our definition of a degenerate two form is equivalent to Hamilton's equations of motion with a zero Hamiltonian:

$$\tilde{\omega}(X, \cdot) = dH_{ex} \tag{5.7}$$

$$= 0 \tag{5.8}$$

The immediate consequence of the degeneracy is that no unique Hamiltonian vector field is defined within the constraint surface and thus that we cannot define a unique Hamiltonian $\mathbb{R}$-action or flow. Correspondingly, our equation of motion (5.8) is only solvable up to an arbitrary factor[27] meaning that the dynamical solutions can only be unparameterised curves in the tangent bundle $\bar{\gamma}_{UPS}$.

The question is then; can we now simply follow a symplectic reduction procedure and then avail ourselves of the standard description of time, change and observable functions? Or does reparameterisation have some unusual feature that necessitates a different approach? To tackle these issues we need to take a closer look at the physical interpretation of both time and its conjugate momentum and in doing so construct a more elegant and general version of reparameterisation invariant mechanics.

5.3 Jacobi's principle and timeless theory

We can associate the time coordinate t (q_0) in extended mechanics with the value taken by a clock external to our mechanical system. In the case of an open system such an interpretation would seem appropriate; but what about if the system is a closed subsystem of the universe? – or even the universe as a whole? In this case there is clearly no physical basis for an external clock and as such we would look to eliminate t as an independent variable. We can do this by the process of *Routhian reduction*[28] which serves to eliminate

[26] This should come as no surprise as this two form must encode the full structure of the constraint and, since this constraint is the Hamiltonian, therefore the dynamics.

[27] This is because (5.8) can be thought of as a linear homogenous equation which only determines the velocities up to a scaling factor applied everywhere along a solution.

[28] A fuller discussion of Routhian reduction in general, and in this case in particular, is given in Lanczos (1970, §5) and Arnold *et al.* (1988, §3.s2).

a cyclic independent variable (i.e., one which only appears in the Lagrangian as a velocity) by using the equations of motion to set its conjugate momentum equal to a constant. Since we have seen above that the conjugate momentum to time is equal to minus the un-extended Hamiltonian of the system we will give the physical interpretation of the constant involved as minus the total energy, E, of the system. Setting the energy as equal to a constant is of course justified for a closed system.

Let us go back to our original Lagrangian formulation of parameterised particle dynamics defined by $L_{ex} : T\mathcal{C} \to \mathbb{R}$

$$L_{ex} = \frac{\bar{T}}{t'} - t'V \tag{5.9}$$

where, as above, $t' = \frac{dt}{d\tau}$ and $\bar{T} = t'^2 T$. By definition the momenta conjugate to t is:

$$p_0 = \frac{\partial L_{ex}}{\partial t'} = -\frac{\bar{T}}{t'^2} - V = -T - V \tag{5.10}$$

Now, since L_{ex} contains t only as a velocity (it is cyclic in t) we can fix the value of p_0 to some constant $c_0 = -E$ by virtue of the relevant Euler-Lagrange equation:

$$\begin{aligned} \frac{\partial}{\partial \tau}\left(\frac{\partial L_{PPD}}{\partial t'}\right) &= \frac{\partial L_{PPD}}{\partial t} = 0 && (5.11)\\ \frac{\partial}{\partial \tau}(p_0) &= 0 && (5.12)\\ p_0 &= -E && (5.13)\end{aligned}$$

This means we can now express t' as a function of $\bar{T}$, V and E:

$$t' = \sqrt{\frac{\bar{T}}{E - V}} \tag{5.14}$$

We can now pass to the Routhian functional:

$$\begin{aligned} R &= L_{ex} + Et' && (5.15)\\ &= \frac{\bar{T}}{t'} - t'V + Et' && (5.16)\\ &= 2\sqrt{\bar{T}(E - V)} && (5.17)\end{aligned}$$

We find that extremizing R leads to the same Euler-Lagrange equations as were defined on our original tangent bundle $T\mathcal{C}_0$ but with τ still playing the role of an arbitrary change parameter:

$$\frac{\partial}{\partial\tau}\left(\frac{\partial R}{\partial q_i'}\right) = \frac{\partial R}{\partial q_i} \tag{5.18}$$

If we take this action and this space as in fact defining our dynamics (with $\bar{T}$ now simply taken to be the kinetic energy and so written as T) then we have achieved a *Routhian reduction* and arrived at a Jacobi type action principle:

$$\begin{aligned} I &= \int d\tau 2\sqrt{(E-V)T} \\ &= \int d\tau L_J(q, q') \end{aligned} \tag{5.19}$$

This action can be understood as defining geodesics in configuration space without making any reference to time or parameterisation. As such it is reparameterisation invariant. If we then proceed to define a function called the lapse as:

$$N = \sqrt{\frac{T}{(E-V)}} \tag{5.20}$$

we can construct conjugate momenta (defined in $T^*\mathcal{C}_0$) according to the simple form

$$p_i = \frac{q_i'}{N} \tag{5.21}$$

The Jacobi Hamiltonian, $H_J : T^*\mathcal{C}_0 \rightarrow \mathbb{R}$ can then be expressed as:

$$H_J = \sum_i p_i . q_i' - L_J = Nh \tag{5.22}$$

where

$$h = \frac{1}{2}\sum_i p_i . p_i + V - E = 0 \tag{5.23}$$

This is again a first class primary constraint. In fact it is the same constraint as was encountered in extended mechanics merely with p_0 replaced by $-E$ and the multiplier t' replaced by N. Thus, reparameterisation invariant theories of mechanics have a Hamiltonian of the form

$$H = Nh \tag{5.24}$$

where N is a arbitrary multiplier, the choice of which determines the parameterisation, and h is some function of the conjugate variables that is equal to zero. Such *timeless* theories will inevitably be constrained Hamiltonian theories with the Hamiltonian itself playing the role of the constraint. Thus the geometry of the constraint surface will be dictated by the two form $\omega = \mathbf{d}\theta = \mathbf{d}(\theta_{|\Sigma})$ where $\Sigma = \{x \in \Gamma : H = 0\}$.

This two form will in general be closed and it will also be degenerate since it has a null direction associated with the Hamiltonian constraint. The integral curves of this vector are the gauge orbits of ω on Σ. However, since this null vector field on the constraint surface is generated by the Hamiltonian we could also argue that $\omega(X) = 0$ is the equation of motion.[29] Since the integral curves of the kernel of the presymplectic form can be shown to be unique solutions we have the strange situation in timeless mechanics where the gauge orbits correspond to the physical histories! The question of how we are to interpret such a perplexing description of mechanics, where degeneracy and dynamics are so closely interwoven, is far from trivial and shall occupy us for much of the remainder of this chapter. To go forward, however, we must go back and reconsider the connection between presymplectic geometry and local symmetry groups.

5.4 Degeneracy, indeterminacy and triviality

In our initial discussion of presymplectic geometry we associated the degeneracy encountered with a group of local or gauge symmetries arising on the tangent bundle to some configuration space, $T\mathcal{C}$. These symmetries were taken to be such that they allow for multiple points to be associated with the same value of the Lagrangian and thus ensured that the Legendre map, $\mathcal{FL} : T\mathcal{C} \rightarrow T^*\mathcal{C}$, was not an isomorphism (a bijective homomorphism) since in such a situation it will generically neither be injective nor surjective. In the case of reparameterisation invariant theory the relevant symmetry group is of course that of reparameterisations. It can be seen to be different to the generic gauge group considered in §1.5 in two important respects. First, since it relates curves that differ in terms of parameterisation it is strictly a symmetry of the action rather than the Lagrangian. Sec-

[29] This can be explicitly seen for the case of the simple pendulum system used by Rovelli (2004) to illustrate both extended mechanics (§3.1 pp. 104-5) and Jacobi's theory (§3.2 pp.109-11) – n.b. he refers to the latter non-standardly as *relativistic* mechanics.

ond, although it also leads to a Legendre transformation that is again not bijective (since it is not injective) the action of the reparameterisation group is such that the conjugate momenta are not affected by rescaling the parameter. Thus, distinct points on the tangent bundle which can be mapped from one to another by the action of the reparameterisation group will correspond to single points on the cotangent bundle. The structure of our phase space is therefore such that paths through it are invariant under reparameterisations. The degeneracy present does not then lead to the type of pernicious underdetermination which was encountered in the construction of presymplectic mechanics considered in §1.5. Rather it takes us between vector fields that are equivalent up to scaling by a multiplicative factor corresponding to the parameterisation. Our primary motivation for the application of the symplectic reduction procedure is therefore removed since there is no possibility of pernicious indeterminism.

We still, however, have the problem of representing change within the presymplectic constraint surface (Σ, ω) – one would like to be able to associate the Hamiltonian with a unique vector field and therefore be able to establish a unique flow with which we can associate evolution. The most obvious way to do this would be to find an underlying symplectic manifold within the timeless theory – thus it may be worth trying to symplectically reduce such theories even without a pressing theoretical need to. However, as pointed out above, timeless theories have a geometry such that what we would normally call the gauge orbits (since they are the sets of points connected by parameterisation rescalings) are also the usual candidates for the solutions in phase space (since they are generated by the Hamiltonian). Thus, the reduction procedure whereby we quotient out the orbits of ω, will leave us with a reduced phase space, $\Pi_R = \Sigma/Ker(\omega)$, without any meaningful notion of evolution – it consists of unconnected points each of which can only gain meaning when referred back to the entire history on the constraint surface to which they correspond. Moreover, since the space is equipped only with a trivial Hamiltonian function there is no sense in which the reduced phase space symplectic form, ω_R, found in reparameterisation invariant theories of mechanics can play any meaningful role – even in generating maps between points in the reduced space. Thus, representationally Π_R alone is only equipped to describe trivial universes consisting of one static configuration (Maudlin (2002) makes a similar point). Furthermore, since ω_R is defined only in virtue of the constraint surface via $\omega = \pi^* \omega_R$ there is a sense in which it could be said to have no more than a purely formal existence.[30]

[30]Rovelli's (Rovelli (2004)) treatment introduces ω_R as ω_{ph} (p.111) but fails to make any use of it.

It could be argued Belot (2007, p.78) in this context that points in the reduced phase space should be taken to describe entire dynamic solutions and therefore that the space is not representationally trivial. In normal circumstances it is reasonable to interpret the reduced phase space, Π_R, resulting from the application of symplectic reduction as a space of instantaneous initial data states, $\mathcal{I}$. This follows from the fact that for any curve γ_{PS} in the space of gauge invariant solutions to the Euler-Lagrange equations $\mathcal{S}_R$ we can define a set of isomorphisms between Π_R and $\mathcal{S}_R$ such that for each value of the curve's parameterisation there will be a map uniquely picking out a point in Π_R with corresponding value of the Hamiltonian flow parameter.[31] However, for the case of nonrelativistic[32] timeless theory there is only a single canonical isomorphism defined between points in the reduced phase space and the *unparameterised* gauge invariant solutions, γ_{UPS}. Thus we can see why one might think the representational role of Π_R should be modified such that it becomes identical to that of $\mathcal{S}_R$. But such a move has highly nontrivial consequences for how we must interpret the *unreduced* phase space and is therefore difficult to countenance. In particular, if $x_R \in \Pi_R$ is a solution then given a point on the constraint manifold in the unreduced phase space, $x \in \Sigma$, we must interpret the relevant 'gauge' orbit, $[x] : \Sigma \to \mathbb{R}$, as an equivalence class of solutions. This interpretation cannot hold since these orbits are equivalent to solutions themselves rather than equivalence classes of solutions. Thus, in nonrelativistic timeless theory at least, the representational role of the reduced phase space cannot be in describing entire histories – we cannot treat it as a primitive arena for representing our fundamental ontology. Rather, any status it can be given as a history space is purely parasitic on the pull-back map to the unreduced space and it is fallacious to argue that the isomorphism that exists between $\mathcal{S}_R$ and Π_R must confer representational equivalence between these two very different mathematical structures.

It would seem therefore that we have established two examples of mechanical theory within which the presence of a first class constraint does not indicate that a symplectic reduction is appropriate. This means that Dirac's theorem (first class constraints generate gauge symmetry) does not hold for the timeless theories considered and is therefore not generally valid in its original form.[33]

[31] The geometric structure of such a reduced space of solutions as well as its connection with the Hamiltonian framework is extensively discussed in Belot (2007)).

[32] In this respect general relativity would seem to be identical to nonrelativistic theory. Belot's argument (which was designed for application to GR) is explicitly re-examined for the case of relativistic theory in §11.3

[33] Rather we should say that first class constraints indicate the presence of gauge symmetries but need not necessarily be identified as the relevant generators. This point is in full agreement with Barbour and Foster

(2008).

Chapter 6

Representing change and observables in timeless mechanics

The essential point established by our argument thus far is that the unreduced phase space of a timeless system (i.e., one in which the Hamiltonian is a constraint) is such that we cannot interpret it using the conventional machinery of constrained Hamiltonian mechanics. Although, as in the generic case, points not on the constraint surface must be classified as inaccessible states, it has been demonstrated that, unlike in the generic case, the difference between points connected by the orbits generated by the constraint on the constraint surface itself cannot be classified as purely unphysical gauge without trivialising the theory. Thus, the geometric structure of timeless theories leads us into an acute *problem of representing change* since we cannot avail ourselves of the conventional temporal machinery provided by a reduced phase space. The definition of a Dirac-Begmann observable also becomes ambiguous within timeless theory since by application of the third condition from §1.5 observable functions can only be equivalent to single points on any dynamical history that is represented on reduced phase space – and this would seem to trivialise them. Furthermore, the first condition (constancy along gauge orbits on the constraint manifold) can only be satisfied in the case of phase space functions which are constant along entire histories of the system and it is difficult to see how such functions – *perennials* in the terminology of Kuchař (1992) – could be used to represent dynamic physical quantities since they cannot change along the solutions defined by the Hamiltonian on the constraint surface. Thus we are also presented with a *problem of representing observables*. This chapter will outline and evaluate two methodologies each designed to meet our two problems for the case of nonrelativistic theory.

6.1 The emergent time strategy

That the Hamiltonian constraint in reparameterisation invariant theories should be thought of as generating genuine change is a position that has been notably defended by Kuchař (Kuchař (1991b, 1992)) and Barbour (Barbour (1994, 2009)); more recently it has been outlined explicitly in Barbour and Foster (2008). We shall call it the Kuchař-Barbour-Foster (KBF) position with regard to change. In keeping with our discussion above, it is an explicitly non-reductive strategy since it involves our treating the differences between points on the integral curves corresponding to the Hamiltonian vector field as genuine physical change. Parallel, although logically independent, to this position with regard to change is the view that observable functions need not commute with the Hamiltonian – we shall call this view the KBF position with regard to observables. This explicitly non-reductive strategy characterises observables as full functions on the unreduced phase space which are allowed to break all three of the Dirac-Bergmann criteria. Essential to the practical viability of this position is the possibility of quantifying the change of an observable in a gauge invariant manner and we shall here outline the methodology for doing this uniquely by using an *emergent* notion of time following Barbour and Foster (2008).

From above we have that a generic timeless Hamiltonian will be of the form:

$$H = Nh \tag{6.1}$$

$$h(p,q) = 0 \tag{6.2}$$

If we take a function on phase space $g(p,q)$ which we would like to interpret as corresponding to some physical quantity then, since the full phase space is a symplectic manifold, we can define the Poisson bracket of this function with the Hamiltonian function, $\{g,H\}$. This is equivalent to the Lie derivative of the function with respect to the Hamiltonian vector field, $\mathcal{L}_{X_H}(g)$. Since the Lie derivative is an operation on scalar functions that gives us the change of the function along a vector field $\mathcal{L}_{X_H}(g)$ is equivalent to a real number representing the rate of change of g along the Hamiltonian vector field with respect to an arbitrary parameter τ:

$$\frac{\delta g}{\delta \tau} = \{g,H\} \tag{6.3}$$

Thus an infinitesimal change in the function along the vector field is equivalent to:

$$\delta g = \delta\tau\{g, H\} \quad (6.4)$$

$$= \delta t\{g, h\} \quad (6.5)$$

where we have introduced the temporal increment $\delta t = N d\tau$. Crucially, we have from the invariance of the canonical action that $N d\tau$ must be invariant under reparameterisations. Since the Poisson bracket must be a real number δg must itself also be a reparameterisation invariant quantity. However, it cannot yet be taken to represent the change in a physical quantity; we have not made any restriction to the constraint surface so we have not excluded change that takes us from accessible to inaccessible states. To resolve this we introduce the weak inequality and the infinitesimal change of a dynamic variable along a physical history can be then represented as:

$$\delta g \approx N d\tau\{g, h\} \quad (6.6)$$

We can put this result in the context of our geometric discussion since we have that: i) The Hamiltonian can be taken to generate an equivalence class of vector fields, X_{Nh} upon phase space[34]; ii) The integral curves of each of the vector fields will correspond to the same set of solutions only with a differently scaled parameter τ marking out change along them; iii) A reparameterisation is then the map between one vector field and another (between one solution and another) by re-scalings of τ. Such a change is between different objects both generated by H but is not strictly generated by H itself. Thus it should come as no surprise that there is a viable methodology for gauge invariantly using the vector fields associated with the unreduced Hamiltonian to solve our problem of representing both change and observables in timeless theory.

Although we now have a valid methodology for representing the change of a function along a timeless solution there does still seem to be a problem. If we were to consider astronomers in two nonidentical isolated sub-systems each using these equations to describe the dynamics of their solar system, they would end up arriving at two different measures of change since each will have to make an arbitrary choice in the form of the lapse and parameter τ. However, if we make the restriction that we are dealing with closed systems of fixed energy then we are justified in fixing the form of the lapse in accordance

[34]We get an equivalence class rather than a unique field because the multiplier N is arbitrary.

with Jacobi's theory – i.e., such that $N = \sqrt{\frac{T}{(E-V)}}$. This Jacobi lapse allows us to define a uniquely distinguished and reparameterisation invariant *Newtonian temporal increment*[35]:

$$\delta t = \sqrt{\frac{T}{(E-V)}}d\tau \tag{6.7}$$

Furthermore, this Newtonian temporal increment is such that it can be defined based purely upon change in the configuration variables as:

$$\delta t = \sqrt{\frac{\delta q_i.\delta q_i}{2(E-V)}} \tag{6.8}$$

and we can therefore represent the change in a function along a solution without reference to the parameterisation. This means that we can treat time as something which naturally *emerges* from the dynamics and is thus ontologically secondary to the change of configuration variables.

6.2 The correlation strategy

An alternative, and perhaps more radical, methodology for representing change and observables in timeless mechanics places emphasis on the idea of correlations and may be traced back through a linage featuring famous names such as DeWitt (1967), Bergmann (1961), and (arguably) Einstein (1916). Here we will present a particular implementation of the correlation strategy which follows on from Rovelli's (Rovelli (1990, 1991, 2002b, 2004)) complete and partial observables methodology and is due to Dittrich (2006, 2007) and Thiemann (2007). We shall focus initially on this correlation strategy as addressing the problem of representing observables in isolation from the problem of representing change and shall designate the position outlined as the Rovelli-Ditterich-Thiemann (RDT) observables position.

An essential element of this scheme is the move away from a representation of change in an observable as the variation of a phase space function along a history. Rather, we focus upon the configuration variables themselves (the partial observables) and assert that the quantities we should be interested in endowing with physical meaning are the relations

[35] As pointed out by Barbour (1994, §4) this privileged time measure derivable from dynamics of a closed system is equivalent to the astronomers notion of ephemeris time.

between configuration variables (the gauge invariant complete observables).

Change in an observable can then be represented as the reparameterisation invariant specification of the value of one configuration variable with respect to another – as correlations between partial observables. The complete observables are the families of correlation functions which individually give the value of one of the partial observables when the other (the clock variable) is equal to some real number.

There is some debate as to how we should interpret the partial observables. In some of his later treatments Rovelli seems to imply that they can be considered to have some independent physical reality – they are 'the quantities with the most direct physical interpretation in the theory' (Rovelli, 2002b, p.124013-7) and, moreover, 'we can associate [them with] a (measuring) procedure leading to a number' (p.124013-2). However, the viability of this interpretation has contested by both Thiemann (2007, p. 78) and Rickles (2008, pp.154-68) principally on the grounds that it is difficult to see how such quantities could be understood as physical magnitudes within a gauge invariant framework. A particular problem is how we can understand a measurement of a single partial observable as possible independent of anything else – surely it makes more sense to view a measurement itself as a complete observables (i.e., a correlation between the values of two quantities). Furthermore, by definition a theory cannot make any *predictions* with regard to partial observables, so it seems difficult to motivate endowing them with any empirical significance. In what follows we shall follow Thiemann, Rickles and early Rovelli in holding that it should only be the complete observables that are taken to be physically meaningful, and associated with possible measurements (see Rovelli (2007) for further discussion of this point indicating that his current view seems to be that although both approaches are consistent there are practical advantages to the later position).

A simple example will illustrate the important elements of the complete observables scheme. We can consider a system described by two configuration variables (partial observables) q_1 and q_2 which together with their conjugate momenta obey a Hamiltonian constraint of the form $H[q_1, q_2, p_1, p_2] = 0$. The phase space, $(q_1, q_2, p_1, p_2) \in \Gamma$, will as usual have a symplectic structure. We can use the relevant symplectic form to define the action of the Hamiltonian vector field on an arbitrary function, $X_H(f) = \omega(X_f, X_H) = \{f, H\}$. The flow, α_H^τ, generated by this vector field can then be defined for every $x \in \Gamma$ and we can see this flow as acting on a phase space function, $\alpha_H^\tau(f)(x)$, such that it takes us along the solutions. For our system therefore we calculate $\alpha_H^\tau(q_1)(q_1, q_2, p_1, p_2)$ and $\alpha_H^\tau(q_2)(q_1, q_2, p_1, p_2)$. We then designate one of our variables as a *clock variable* and seek to invert an expression of the form $T_x(\tau) = \alpha_H^\tau(q_1)(x)$ such that solving $T_x(\tau) = s$ for $s \in \mathbb{R}$ will give us an expression for τ in terms of s and q_1. In general this inversion will only be possible for a specific interval – thus the clock variables are typically going to be at best locally well defined and so are unlikely to be continuous on phase space and this means that the scheme will be difficult to implement in practice. We can then insert the inverted expression into the second flow equation $\alpha_H^\tau(q_2)(x)$ by substituting for τ, and produce an expression which (within the interval specified) gives us the value of q_2 when q_1 takes the value s. This *complete observable* represents a family of functions (one for each s) each of which expresses the correlation between our two partial observables without reference to parameterisation.

Let us explicitly calculate the expression for such a complete observable given a toy model double pendulum system with Hamiltonian constraint of the form:

$$H[q_1, q_2, p_1, p_2] = \frac{1}{2}(q_1^2 + q_2^2 + p_1^2 + p_2^2) - E = 0$$

The Dittrich methodology involves taking the configuration variables and explicitly constructing the flow generated by the action of the constraint on each of these variables. For our system therefore we are looking to calculate $\alpha_H^\tau(q_1)(q_1, q_2, p_1, p_2)$ and $\alpha_H^\tau(q_2)(q_1, q_2, p_1, p_2)$. To do this we expand the action of the flow in terms of a power series:

$$\alpha_H^\tau(f)(x) := e^{\tau \mathcal{L}_{X_H}}(f) = \sum_r^\infty \frac{\tau^r}{r!}\{H, f\}_r(x)$$

where $\{g, f\}_{(0)} := f$, $\{f, g\}_{(n+1)} := \{g, \{g, f\}_{(n)}\}$ is the iterated Poisson bracket. For our Hamiltonian this gives:

$$\begin{aligned}
\alpha_H^\tau(q_1)(q_1, q_2, p_1, p_2) &= q_1 \cos(\tau) - p_1 \sin(\tau) \\
&= \sqrt{q_1^2 + p_1^2} \sin\left(\tau - \arctan(\frac{q_1}{p_1})\right) \\
\alpha_H^\tau(q_2)(q_1, q_2, p_1, p_2) &= q_2 \cos(\tau) - p_2 \sin(\tau) \\
&= \sqrt{q_2^2 + p_2^2} \sin\left(\tau - \arctan(\frac{q_2}{p_2})\right)
\end{aligned}$$

As above we then: i) Designate one of our variables as a *clock variable* $T_x(\tau) = \alpha_H^\tau(T)(x)$; ii) Seek to invert the expression so that solving an equation of the form $T_x(\tau) = s$ for $s \in \mathbb{R}$ will give us an expression for τ in terms of s and the configuration variables; iii)

Insert the inverted expression into the second flow equation, $f_x(\tau) = \alpha_H^\tau(T)(x)$, and produce an expression for the value of one partial observable when the other takes the value s:

$$F_{[f,T]}(s,x) = \alpha_H^\tau(f)(x)_{|\alpha_H^\tau(T)(x)=s}$$

Designating q_1 as the clock variable and focusing on the interval $[-\frac{\pi}{2} + \arctan(\frac{q_1}{p_1}), \frac{\pi}{2} + \arctan(\frac{q_1}{p_1})]$ leads us to invert $\alpha_H^\tau(q_1)(x) = T_x(\tau) = s$ to get:

$$\tau = \arcsin\left(\frac{s}{\sqrt{q_1^2+p_1^2}}\right) + \arctan(\frac{q_1}{p_1})$$

Inserting this into $\alpha_H^\tau(q_2)(x)$ gives:

$$\begin{aligned} F_{[q_2,q_1]}(s,x) &= \alpha_H^\tau(q_2)(x)_{|\alpha_H^\tau(q_1)(x)=s} \\ &= \sqrt{q_2^2+p_2^2}\sin\left(\arcsin\left(\frac{s}{\sqrt{q_1^2+p_1^2}}\right) + \arctan\frac{q_1}{p_1} - \arctan\frac{q_2}{p_2}\right) \end{aligned}$$

For the interval specified, this expression gives us the value of q_2 when q_1 takes the value s. As such this complete observable represents a family of functions each of which expresses the correlation between our two partial observables without reference to parameterisation. A more general expression can be defined based on translations from our original interval by $k\pi$ and $2k\pi$ for $k \in \mathbb{Z}$ – see Dittrich (2007, p.1898). If we then project down into the configuration space $(q_2, q_1) \in \mathcal{C}$ using the map $P : \Gamma \to \mathcal{C}$ given by $x \mapsto (q_1(x), q_2(x))$ for $x \in \Gamma$ then the phase space flows will map into flows in $\mathcal{C}$. The image of these flows define unparameterised curves in configuration space and for our example these can be shown to be ellipses. This explicitly demonstrates correspondence between Dittrich's methodology and that originally applied to this system by Rovelli (1990).

Both in this specific case, and in general, we can see that not only are complete observables families of reparameterisation invariant objects, but the functions on phase space that each correlation defines will commute with the Hamiltonian constraint. This means that they explicitly fulfil the second condition for a Dirac-Bergmann observable and demonstrates the fundamental difference between the RDT and KBF positions with regard to observables. We can consider the extent to which the complete observables satisfy the other two criteria. The first condition was that Dirac-Bergmann observables are functions which are constant along the orbits generated by the constraint on the constraint surface. By definition the flows generated by the Hamiltonian constraint in the phase space and the integral curves of the relevant null vector field will coincide on the constraint surface. Since each of the correlations that make up a complete observable are defined for a specific value of the flow parameter these functions do not vary along this flow and are therefore constant along gauge orbits. But it must be noted that the sense in which these functions satisfy this condition is somewhat different from the generic case in two senses. First, in a typical gauge theory an observable would be constant along gauge orbits but it would also vary between them – it is this variation off the orbits that we would normally consider physical change. Second, the sense in which they are constant on gauge orbits is almost trivial – they are each defined for a particular value of the flow parameter so in effect they establish the correlation at a particular point along an orbit. Clearly such a specification is valid all the way along the orbit only in the same strange sense that 'in Sydney in 2011 AD, Caesar crossed the Rubicon in 49 BC' is a valid statement concerning modern Australian history.

Application of the third Dirac-Bergmann condition is more acutely problematic. Since *for a given dynamical solution* the functions that define the observables cannot, by definition, vary between gauge orbits, the complete observables relevant to an individual solution are each equivalent to single points rather than functions on the reduced phase space. This means that if we take the symplectic reduction ontologically seriously (i.e., treat the reduced phase space as primitive) we will, for any given dynamical solution, only be left with a single correlation specified by each complete observable rather than an entire family of correlation functions since it is only through the pull back to the constraint manifold that these correlations are defined. It would seem, therefore, that there is some motivation for setting aside the Dirac-Bergmann notion of an observable altogether – complete observables are defined in such a way that it is no longer fully appropriate and the RDT position should be seen as a distinct alternative rather than a innovative application of the orthodoxy.

We can now finally turn the problem of change. Here we appear to have a problem since Rovelli and Dittrich hold both that evolution generated by the Hamiltonian is gauge[36] *and* that the entire orbit it generates is what should be considered physically real.[37] If we dispense with the first proposition (which clearly must contradict the non-

[36]See Rovelli (2004, p.127) and Dittrich (2007, p.1892). Thiemann's Thiemann (2007, p.75) position with regard to this point is more nuanced and is specifically targeted to the case of general relativity.
[37]See Rickles (2008, pp.182-6) and Dittrich (2007, p.1894).

reductive stance taken by these authors) and focus on the second, then a coherent but highly radical position emerges. In particular, if we consider the implications of the change in the notion of the physical state that seems to have been made, then it appears that the RDT position with regard to change in nonrelativistic reparameterisation invariant mechanics amounts to a denial of the need for any fundamental concept of time at all.

Rovelli (2002b) distinguishes the 'physical phase space' as the 'space of orbits generated by the constraints on the constraint surface' (p3) and Dittrich (2007) similarly defines the physical state as an 'equivalence class of phase space points' which 'can be identified with an n-dimensional gauge orbit' (p 1894). For a theory where the Hamiltonian is itself a constraint this constitutes a redefinition of the structure of our dynamics such that the basic ontological entity is an entire history rather than an instantaneous configuration. In typical gauge theories points on the constraint surface connected by a gauge orbit are classified as the same state because the difference between them is taken to be unphysical – we can then proceed to a symplectically reduced phase space within which we can characterise the change between two instantaneous states without problem. This interpretation of change drawn from the complete observables scheme on the other hand leads us to classify two such points as the same state because the word 'state' is redefined such that in includes all points on the orbit. This is not to classify time or evolution as gauge since that would indicate that the trivial reduced phase space of single initial data points was the arena of true physical significance. Rather, it is to adopt a position such that any notions of evolution and time in a conventional sense are redundant within reparameterisation invariant theory. Adoption of a correlation strategy has then the capacity for radical philosophical implications for the nature of time in physical theory – the next section will examine these in more detail as well as considering the emergent time strategy in a more philosophical context.

Chapter 7

Interpretational implications

The objective of Chapter 5 was to demonstrate that, unlike standard gauge theories, timeless nonrelativistic theories are such that the constraints cannot be considered as gauge generators without trivialisation and that a reduced phase space with a symplectic geometry cannot be considered as both a viable and autonomous representative structure. In Chapter 6 we examined two strategies for representing observables and change in the unreduced phase space and considered some of the implications of each scheme. What now concerns us are the interpretational consequences we should attach to our conclusions. In particular, it is interesting to consider how we should place the existence of: 1) gauge theories with phase spaces such that passage to a representatively viable reduced space is not available; and 2) our two strategies for representing change without an explicit notion of time; in the context of the debates over both relationalism/substantivalism with respect to time and reductionism/non-reductionism with respect to the interpretation of gauge theories.

7.1 The relationalist vs substantivalist dispute with regard to time

The long standing relationalist/substantivalist dispute with regard to space and motion in nonrelativistic mechanics contains many important lessons for the parallel dispute with regard to time. In particular, modern treatments in terms of analytical mechanics allow us to characterise precisely a number of refinements to the traditional binary distinction – we will very briefly introduce the ideas key for our purpose, a more exhaustive analysis can be found in Rickles (2008).

Let us define a *substantivalist* as someone who is committed to the existence of space (or space-time) as an entity in its own right, over and above the relations that hold between material bodies. The position of *straightforward substantivalism* then involves a commitment to the existence as distinct possibilities, spatial (or space-time) models which differ only by the application of an element of the Euclidean (Galilean) group of global symmetry transformations. The difference between the two models is naturally cashed out in *Haecceitistic* terms since it rests upon the non-qualitative cross-identification between spatial points as the means of differentiation between the distinct possibilities in question. In terms of the Hamiltonian formulation of mechanics (where the models are represented by curves in phase space) straightforward substantivalism involves insisting that sequences of points in phase space which are related by symmetry transformations can represent distinct sequences of instantaneous states since they differ Hacceistically in terms of the differing roles that spatial points play between gauge related instantaneous states.

A *sophisticated substantivalist* is someone who maintains the commitment to the ontological fundamentality of space (or space-time) but insists that models related by Euclidean (Galilean) symmetry transformations are not distinct possibilities. The most natural way of securing this reduction in possibilities is to adopt a position of anti-Hacceitism such that we do not allow non-qualitative determinants of cross-identification, and therefore differences between spatial models solely as to which spatial points play which roles are not allowed: models that differ solely Haecceitisticly are counted as the same possibility. There is a natural implementation of sophisticated substantivalism within Hamiltonian mechanics grounded upon anti-Hacceitism: one can insist that sequences of points in phase space which are related by symmetry transformations cannot represent possibilities since they refer to sequences of instantaneous states which differ solely as regard to which spatial points play which roles.

The *relationalist* on the other hand, wants to deny that space (space-time) is a fundamental entity and is therefore committed to denying that models which differ only with regard to space (space-time) symmetries constitute distinct possibilities. They can ground this reduction in possibilities *without* having to endorse anti-Haecceitism, since because they deny that there are spatial (space-time) points, they automatically have that models related by Euclidean (Galilean) symmetry transformations constitute the same possibility since there simply are no individuals to ground even a Haecceitistic difference. Space (space-time) relationalism can be naturally implemented within a Hamiltonian system of mechanics either by focusing upon the original phase space and identifying as the same possibility instantaneous states which differ solely with regard to the application of a Euclidean (Galilean) symmetry transformation, or by moving to a quotient space where all points related by elements of the relevant symmetry group are reduced to single points. The latter approach amounts to recovering phase space literalism within a reduced phase space and, for reasons that shall become important later, will be considered as a distinct position of *reductive relationalism* (like reductionism with respect to gauge theory, it is notably advocated by Belot (1999, 2000)). There is of course nothing to preclude a sophisticated substantivalist from also adopting the formalism of the reduced phase space since they count possibilities in an identical fashion.

With these distinctions in hand, and the existence of a connection between reductionism and relationalism already apparent, we can turn our attention to the ontological status of time within our timeless theories of nonrelativistic mechanics. We can define a temporal substantivalist as someone who asserts the existence of time as a basic entity in its own right over and above the relations that exist between the instantaneous states of material systems (be they relationally defined or not). Such a position is a natural reformulation of the Newtonian concept of absolute time; in particular, it seems to implement that notion of time defined in the influential Scholium section of his Principia.[38] Now, it could be argued that, at least as nonrelativistic mechanics is concerned, substantivalist time is inherently connected to the use of an external temporal dimension and on this basis a substantivalist would have a very hard time dealing with Jacobi's theory. However, what is essential to temporal substantivalism – under our reading of it at least – is that time can be asserted as a basic entity parameterising change that is not parasitic on the motion of the bodies that are doing the changing. Thus, Jacobi's theory does not in principle exclude temporal substantivalism since change is parameterised (albeit non-uniquely) in terms of τ. Moreover, unlike its Newtonian counterpart (as well as parameterised particle mechanics) Jacobi's theory offers a level playing field for matching the temporal substantivalist against their relationalist foe since it is a mechanical framework free from the fundamental *presumption* of preferred parameterisation or external time that would inherently favour a substantivalist reading.

A straightforward (i.e., Haecceitist) temporal substantivalist reading of Jacobi's theory could then proceed as follows. Just as the reality of space indicates that there is a real but non-qualitative difference between two sequences of instantaneous states related by a spatial symmetry transformation, the reality of time indicates that there is a real but non-

[38] 'Absolute, true, mathematical time, of itself, and from its own nature, flows equably without relation to anything external, and by another name is called duration: relative, common time, is something sensible and external (whether accurate or unequal) measure of duration by which the means of motion, which is commonly used instead of true time' Newton (1962).

qualitative difference between two sequences of instantaneous states related by a temporal symmetry transformation. In the first case this difference is represented by sequences of points in velocity-configuration/phase space differing only with regard to the application on an element of the Galilei group of global space-time symmetries. In the second it is represented by two sequences of points in velocity-configuration/phase space differing only with regard to an application of an element of the reparameterisation group. In each case this non-qualitative difference can be understood precisely in Haecceitistic terms because it is established via inter-structure cross-identification of individual instantaneous states (they play different roles in the different structures). That these models are connected by an element of the local symmetry group of time reparameterisations does not mean that they fail to be distinct because, even though such a symmetry means that there can be no empirical difference between worlds which differ only with respect to their parameterisation, our acceptance of Haecceitism allows us to say that there is an ontological difference. Thus, the straightforward substantivalist type position with respect to time in Jacobi's theory leads us to endow parameterisation of solutions with a stamp of physical reality. We can thus see straightforward temporal substantivalism as a direct application of the histories Hacceitism introduced in Chapter 2.

Correspondingly, Jacobi's theory, at least as formulated in §5.3, leaves open the conceptual space for a sophisticated (i.e., histories anti-Haecceitist) form of temporal substantivalism whereby time is still asserted as a basic ontological entity but the fundamental temporal structure of a sequence of instantaneous states is multiply realised in terms of the different parameterisations of a solution – a single fundamental notion of time is understood as being represented by the equivalence class of parameterisations. We do not have an inflation of possibility within the representation of histories since the difference between two parameterisations of a solution is understood to be merely of the excluded histories Haecceitist variety (it is only which instantaneous states play which roles that is different).

A temporal relationalist can be defined as someone who treats time as a non-fundamental or derived entity. Such an anti-Newtonian position is typically seen to have originated with the work of Descartes, Leibniz and perhaps also Huygens (Barbour (2009)). but is contained in the most direct form within the ideas of Mach.[39]. Here we will characterise temporal relationalism as an ontological position such that the basic ontology excludes

[39] 'It is utterly beyond our power to measure the changes of things by time. Quite on the contrary, time is an abstraction, at which we arise by means of the change of things' Mach (1960).

temporal structure beyond an ordering of instantaneous states.

(n.b. Such a definition is currently in full accordance with the more minimal notion of a 'Leibnizian relationalist' with respect to time (Pooley and Brown (2001)). However, its Machian aspect will be further developed as a more positive position of *Machian temporal relationalism* during our discussion of relational clocks in §9.2 as well as during the relativistic treatment in §12.1. In essence the emergent time strategy is a realisation of exactly the type of Machian temporal relationalism that shall be considered later.)

With regard to Jacobi's theory temporal relationalism should be understood as an insistence that the parameterisation of a solution is non-fundamental since the temporal separation between two instantaneous states is excluded from the basic ontology and thus parameterisation – which represents this separation – must be treated as merely an abstraction.

Just as the spatial relationalist was committed to two points in either the velocity-configuration space or phase space which are connected by spatial symmetries representing the same possiblity, the temporal relationalist is committed to two parameterisations of a solution within the relevant space representing the same thing. This would seem, *prima facie*, to leave open the option for either reductive variant of temporal relationalism whereby we quotient out the relevant symmetry group to leave a reduced space with the requisite reduced set of possibilities are represented literally.

So far the debate seems to resemble closely that for space/space-time. However there are two new and interesting complications that we must consider. The first stems from the fact that the reparameterisation symmetry of Jacobi's theory is, unlike the global symmetries that feature in the space/spacetime debate, manifestly local. The locality of the symmetry means that a straightforward substantivalist who sticks with Haecceity and an unreduced possibility space could be left open to pernicious indeterminism in their ontology of the type discussed in chapter 2 when we considered the generic variant of histories Hacceistim. Such a development has been key to the perceived derailment of straightforward substantivalism for the case of general relativity which features local space-time symmetries[40] and may be expected for this case also. Our straightforward temporal substantivalist is understood to be committed to (histories) Haecceitism in that that they admit cross-identification of temporally relabelled instantaneous states between histories as represented by curves related by reparameterisation. Thus, the differently parameterised

[40] This is in fact the essence of the hole argument, see Rickles (2008, Chapters 4-5) for a more extensive discussion.

curves are taken to represent ontologically distinct structures. Such an ontological distinction between objects differing by the application of the action of a local symmetry group has the potential to generate ontological indeterminism since the two curves may initially coincide and then diverge. Since Jacobi's theory is an empirically deterministic theory this potential for ontological indeterminism seems unattractive and could be taken to drive us away from the straightforward variant of temporal substantivalism on the grounds of the commitment to Haecceitism involved.

However, the case of Jacobi's theory is particularly interesting because although pernicious (ontological) indeterminism *is* possible within the velocity-configuration space of Jacobi's theory – since the velocities are dependent on parameterisation – it *is not* possible within the phase space since reparameterisations are symmetries on the canonical momenta. This means that provided they confine themselves to the constraint manifold, a temporal substantivalist can stick to a completely literal reading of phase space built upon histories Haecceitism – i.e., such that each point represents a distinct instantaneous state and each solution representing a distinct dynamical history – without the possibly of ontological underdetermination.

Explicitly we have that: On the one one hand, within a generic gauge theory's phase space along with histories Haccesitism, inevitably goes the possibility of ontological indeterminism – an initial specification of an ontology may have multiple possible continuations corresponding to different gauges. On the other hand, within Jacobi's theory's phase space one can adopt straightforward substantivalism and therefore (histories) Haccesitism without such a problem – an initial specification of an ontology (which includes parameterisation) always provides for a deterministic continuation. Thus even though Jacobi's theory can be classified as a gauge theory in that it features first class constraints, it has a phase space that can unproblematically accommodate an ontological deterministic, non-reductive interpretation without any recourse to anti-Haecceitism or relationalism. In this respect it constitutes a notable counter-example to accounts of the interpretation of gauge theories (such as that presented by Belot and Earman (2001)) which are presumed by their authors to hold generically.

The second point that marks the substantivalism vs relationalism dispute with regard to time in Jacobi's theory distinct from both the case of global symmetries in Newtonian mechanics and local symmetries in generic gauge theories is that the reductionist position is no longer available. As discussed extensively above, the structure of Jacobi's theory is such that the application of symplectic reduction will lead to a reduced phase space which has a trivial dynamical structure such that it can only be made sense of by reference back to the unreduced space. This renders a reductionist reading of the theory inadequate since to get off the ground it would require the utilisation of exactly the otiose structure (*gauge* related points on the constraint manifold) the elimination of which was its supposed benefit. Moreover, the reductionist desire to construct a reduced phase space which can be interpreted along literal lines manifestly fails since on its own the relevant reduced space can only be read as representing isolated instantaneous states corresponding to dynamically trivial universes. Thus, with regard to time in Jacobi's theory at least, any viable form of relationalism is going to have to be non-reductive. Let us then consider the relationalist credentials of our two non-reductive strategies for representing change and observables.

7.2 An ontology of timeless change?

As discussed above the emergent time strategy explicitly makes use of the Hamiltonian constraint as the generator of evolution. A point on the constraint manifold is taken to represent an instantaneous state and the dynamical change between this state and the next is represented in terms of the null vector corresponding to the flow generated by the Hamiltonian at that point. Similarly, an observable is represented by a function of the constraint manifold and the change in an observable is represented by the change in that function along the Hamiltonian flow. Now, it has been argued by Belot and Earman (1999, 2001) that for the case of general relativity treating the relevant Hamiltonian constraint in such a manner (in particular allowing for observables that do not commute with the Hamiltonian constraint) is the hallmark of a *Heraclitean* position that asserts the fundamentality of time within the theory. Conversely, according to this viewpoint, there is an equivalence between treating the Hamiltonian constraint as gauge generating (and therefore implementing the Dirac-Bergmann criteria for observability) and relationalism. Clearly, adopting such a classification scheme for Jacobi's theory would seem to suggest that we should think about the emergent time strategy in terms of temporal substantivalism.

Pooley (2001) argues that we should adjust this classification scheme such that how we treat the relevant constraints of general relativity is now thought of as a guide to deciding between 'straightforward substantivalism on the one hand and the disjunctive set of sophisticated substantivalism and anti-substantialism relationalism on the other' (p. 15). Thus, under Pooley's scheme the emergent time strategy for understanding change in Jaocbi's theory would be classed as a straightforward substantivalist one with respect to time. However, as has been argued for the case of general relativity Rickles (2008, p. 170) the assertion of such definite connections between the treatment of the observables/Hamiltonian constraint and substantivalist/relationalist distinctions is not in fact justified. There is more potential for metaphysical underdetermination within the formalism than would appear at first sight.

The crucial factor informing Pooley's distinction is the reduction in possibility entailed by how we interpret objects within structures connected by the relevant symmetry. For the case of Jacobi's theory – and actually also in GR itself (§10.3) – this turns on how we understand solutions related by the relevant gauge symmetry and not points connected by the action of the Hamiltonian constraint. In Jacobi's theory one can happily avoid straightforward substantivalism whist still denying that the Hamiltonian constraint generates gauge so long as one describes the *change* of observables (which themselves may fail to respect the Dirac-Bergmann criteria) without reference to parameterisation – it is change in parameterisation that we want to call unphysical not the change that is parameterised! The emergent time strategy is temporally relational since it has removed fundamental temporal structure altogether and allows us to describe change, both of observables and states, without reference to parameterisations. Moreover, since within a Hamiltonian formalism it can make use of a one-to-one representational relationship between points and instantaneous states, on the one hand, and solutions uniquely parameterised via the Newtonian temporal increment and dynamical histories, on the other it can be understood to proceed via an entirely literal interpretation of the phase space. As such it is in fact an irresistibly temporally relational mechanical framework since there is simply no temporal entity available for the substantivalist to reify – in effect a reduction of the possibilities entailed by the multiplicity of parameterisations has been enacted. However, this *reduction* is done by use of the Newtonian temporal increment rather than by a direct geometric reduction of the relevant symmetry.

The correlation strategy is distinguished by providing a reparameterisation invariant description of the change of observables which satisfies the second Dirac-Begrmann criterion of commuting with the constraints but does not make explicit recourse to the reduced space *à la* reductionism. However, as discussed at the end of the last section it leads us to a notion of change which constitutes a radical departure from that used in conventional physical theory. The notion of an instantaneous state is dispensed with and the observables are smeared non-locally along an entire solution as constituted by the *gauge* orbit of the Hamiltonian constraint on the constraint surface in phase space. The fundamentally original manoeuvre is to redefine the idea of a state such that it is closer to the idea of a history than its original meaning. How should we see the correlation scheme in the context of our various forms of relationalism and substantivalism? Clearly it cannot be interpreted in terms of temporally substantivalist ontology since time or even change in the traditional sense do not feature in the relevant formalism. Furthermore, it does not fit naturally into the relationalist picture, as we have defined it, since there is not even an order sequence of instantaneous states within its basic ontology.

Rather, we must consider the possibility that the correlation strategy cannot be naturally interpreted in terms of either a relationalist or substantivalist ontology. If we take the issue of primacy between temporal structure and the relations between instantaneous states of a material system to demarcate the distinction between temporal relationalism and substantivalism then clearly a theory in which there are no instantaneous states or temporal structure will transcend our system of classification. If we define temporal relationalism to mean simply ‘not temporally substantivalist’ then we can happily think of the correlation scheme as relationalist – but if we are to think more constructively about temporal relationalism in terms of its Machian philosophical underpinnings with the concept of time parasitic on relational change, then the correlation scheme is certainly not relationalist with regard to time since even a derived, relational notion of time or instantaneous state cannot be found within the formalism.

One must note here that if one adopts the later Rovelli’s interpretation of the correlation strategy (i.e., that of Rovelli (2002b)), in which the partial observables are treated as physical magnitudes, then it could be argued to be appropriate to understand the scheme as a species of temporal relationalism – or perhaps even (following Belot and Earman’s categorisation scheme) substantivalism (Rickles, 2008, p.165). However, as discussed in at the start of §6.2, there are good reasons to move away from such an interpretation, and rather see only the complete observables as fundamental. In these circumstances, it seems clear that it is inappropriate to understand the complete observables scheme as relational since, not only is there no methodology to derive the relevant sequence of (ordered) instantaneous states available (the reasons for this will become explicit within the discussion of Chapter 9), but the fundamental entities defined within the scheme are by definition temporally non-local. The temporal non-locality of the complete observables means that there is no scope for a recovery of an ontology which *changes* in a substantive sense, over and above the idea of locally parameterising the elements of our ontology in terms of families.

So, what kind of ontology should we give to the correlation scheme then if not a temporal relationalist one? The most obvious option would be to take a starkly Parmenidean one – time is purely an illusion and not even a derived or emergent phenomena. There is no change or evolution, merely correlations and timeless states corresponding to histories which cannot be temporally decomposed into instants. In the context of nonrelativistic mechanics adopting such a radical notion of timelessness would seem undesirable given the viability of other options and this, together with the issue of practical applicability, would seem to push us away from adopting the correlation strategy. For addressing the problem of representing change and observables in nonrelativistic timeless mechanics the emergent time strategy clearly provides us with a better option since its interpretation consequences are far more palatable. The case of general relativity, however, is another matter, and in that arena radical timelessness may become a necessity. Since a number of complications within this more powerful theory must be considered in detail before our arguments can be reconstructed, we will defer this discussion to Part III

Chapter 8

Time in ordinary quantum mechanics

Within a non-relativistic, non-gauge quantum system reached via canonical quantisation, the evolution of a quantum observable can be constructed in terms of the commutator between an arbitrary observable $\hat{A} \in \mathcal{A}$ and the quantum Hamiltonian function $\hat{H}$ (which is also a quantum observable):

$$i\hbar \frac{d\hat{A}}{dt} = [\hat{A}, \hat{H}] \tag{8.1}$$

Alternatively, thinking in terms of the Schrödinger picture we can consider the evolution of a state vector $\psi \in \mathcal{H}$ (an element of the Hilbert space) in terms of the time dependent Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t}\psi = \hat{H}\psi \tag{8.2}$$

Assuming the Hamiltonian is not a function of time leads us to a wavefunction that takes the form $\psi = \varphi(q)e^{\frac{-iEt}{\hbar}}$ where $\varphi(q)$ is the solution to the time independent Schrödinger equation:

$$\hat{H}\varphi = E\varphi \tag{8.3}$$

In either picture we have the basic dynamics represented in terms of a unitary transformation which evolves the states/observables in terms of the unique Newtonian background time that has been inherited from the classical formalism. However, the equations above are not the full story. Depending upon the interpretation of quantum mechanics which one takes, there is scope for further, non-unitary evolution. For example, in the Copenhagen interpretation the reduction of the state vector upon *measurement* surely must be a process that takes place in time and if so, since it can evolve superpositions into pure states, it must be non-unitary. More concretely dynamical collapse models explicitly supplement unitary time evolution with some form of non-unitary decay process. We will not consider at all notions of time related to these additional interpretational structures –

to do so properly would involve entering into a detailed discussion of the measurement problem and this would stretch the already wide remit of this project. Rather our investigation of time in quantum mechanics will be restricted entirely to temporal notions within the basic formalism and not within any particular interpretation.

8.1 Ideal quantum clocks and internal time

What is particularly important for the understanding of time in conventional quantum mechanics is the extent to which we can define and utilise an internal notion of time within the formalism. We can define a quantum clock system in terms of a system which features canonical clock variables $\hat{\tau}_i$ which are such that they, together with their canonical momenta $\hat{\eta}_i$, satisfy the commutation relations:

$$i\hbar\hat{I}\delta_{ij} = [\hat{\tau}_i, \hat{\eta}_j] \tag{8.4}$$

$$0 = [\hat{\tau}_i, \hat{\tau}_j] = [\hat{\eta}_i, \hat{\eta}_j] \tag{8.5}$$

$$i\hbar\hat{I} = [\hat{\tau}_i, \hat{H}] \tag{8.6}$$

where $\hat{I}$ is just the identity operator. The first two expressions merely assert that the $(\hat{\tau}_i,\hat{\eta}_i)$ are canonical variables. More significant is the third equation which enforces the clock variables must be linear functions of the external time parameter since by (8.1) we have that:

$$\frac{d\hat{\tau}_i}{dt} = \hat{I} \tag{8.7}$$

and therefore that

$$\hat{\tau}_i = \hat{\tau}_i^0 + \hat{I}t \tag{8.8}$$

where $\hat{\tau}_i^0$ is an operator defining the zero of the clock.

In practice we can find a number of systems with the requisite properties to be quantum clocks. The fundamental choice is between using linear and cyclic system with angles playing the part of the clock variables in the latter (for more details see Hilgevoord (2005)). Importantly, it is found that in concrete constructions the physical viability of a quantum clock systems depends upon us considering it as isolated – otherwise the energy that can be taken out of the clock system is unbounded since the energy spectrum of an ideal quantum clock ranges from $-\infty$ to $+\infty$. Thus, although a ideal quantum clock provides us with an notion of time that is based upon the motion of a system it still relies upon some degree of externality. We divide out part of the world, call it the clock system and then treat it as external from everything else. Such a notion of time is undemocratic in the sense that it is the motion of a particular privileged system that defines time rather that the collective motion of all the bodies in the universe. Nevertheless it does provide us with a representation of time in quantum mechanics that can be associated with Hermitian operators on a Hilbert space and this is key to understanding the energy/time uncertainty relations to which we now turn.

8.2 Time/energy uncertainly relations

We know that *classically* the Poisson bracket between the position and momenta variables is just $\{q_i, p_j\} = \delta_{ij}$. Quantisation involves the substitution $\{,\} \to i\hbar[,]$ so we now have the commutation relation between the canonical variables now represented as operators :

$$[\hat{q}_i, \hat{p}_j] = i\hbar\hat{I}\delta_{ij} \tag{8.9}$$

Textbook quantum mechanics leads from this relation to the famous position/momentum uncertainty relation:

$$\Delta_\psi q_i \Delta_\psi p_i \geq \frac{1}{2}\hbar \tag{8.10}$$

where Δ_ψ stands for the standard deviation of the relevant expectation value relative to a state vector ψ. Since the derivation is insensitive to which Hermitian operators we are using we have that *for any pair of quantum canonical variables with continuous spectrums a version of the uncertainty relation will hold.* This of course includes our clock variables and their canonical conjugates that we defined above (at least in the linear clock case). So for them we also have that:

$$\Delta\hat{\tau}_i \Delta\hat{\eta}_i \geq \frac{1}{2}\hbar \tag{8.11}$$

Significantly, we also have that $i\hbar\hat{I} = [\hat{\tau}_i, \hat{H}]$ which means (again only for the linear clock, there are some technical complications in the cyclic case) we can derive

$$\Delta\hat{\tau}_i \Delta\hat{H} \geq \frac{1}{2}\hbar \tag{8.12}$$

which a time/energy uncertainty relation since the expectation value of the Hamiltonian is an energy. Of course the τ here is *not* external time t. It is the internal time measured by our ideal quantum clock. In fact, contrary to many treatments there is no implied uncertainty as to the measurement of any observable (including energy) in terms of the external time parameter within the quantum formalism. External time within quantum mechanics is not represented in terms of an operator (again see Hilgevoord (2005)) and observable quantities can be determined with respect to it to an arbitrary degree of accuracy.

8.3 Space vs. time vs. spacetime

The final few sentences of the last section could be taken to indicate a deficiency within quantum theory. In a sense we do have operators and uncertainty relations relating to space in terms of the position operators $\hat{q}$ and their relationship with the momentum operators. Thus, it may seem that quantum theory introduces an asymmetry between space and time that is not only conceptually unsatisfactory but also contrary to the basic principles underlying special relativity.

However, as pointed out by Hilgevoord (2005), we must be careful to distinguish between the position operators $\hat{q}$ and the spatial coordinates. Classically, the q refer to the positions of point particles which are classical observables and therefore represented by operators quantum mechanically. On the other hand spatial coordinates x are not strictly classical observables and therefore quantum mechanically too they are not observables. Of course we can use the q to define an internal spatial reference system and therefore can quantum mechanically consider a internal notion of space in quantum mechanics as defined via operators. However, this is simply analogous to the internal representation of time provided by quantum clocks as discussed above.

Quantum mechanics deals with space and time in an equitable manner; although concepts can be represented internally in terms of operators, the external notions are not quantised. Instead we have a classical spacetime background which provides an un-quantised reference system for the quantum dynamics. Making the theory consistent with special relativity would not involve the quantisation of time or space.

Rather the motivation for eliminating external time from quantum mechanics come from Machian arguments towards the type of temporal structure a theory of mechanics should have. In particular, we will see (in §9.2) that a desire for an equitable, purely internal notion of time drives us towards attempting to construct a quantum analogue to the classical Jacobi type theory that we have been investigating. However, as we shall see in the next chapter, the task of quantising Jacobi's theory such that we can retain the full classical theory in the appropriate limit and construct dynamics with respect to an internal, relational clock is highly non-trivial. In order to accomplish it we will need to construct new techniques.

Chapter 9

Relational quantisation and the quantum problem of time

The logical starting point of this chapter – which chiefly gives an account of ideas developed in Gryb and Thébault (2012)[41] – has been established already within our current discussion (in particular §5.4). The reader should, by now, be convinced that the physical origin of the Hamiltonian constraints that feature in globally reparameterisation invariant theories is such that their phase space action *must* be treated as a physical transformation. In fact, we have seen that they generate precisely the transformations associated with time evolution. In this context, usual gauge theory methodologies become inapplicable: since the integral curves of the null directions associated with Hamiltonian constraints are solutions rather than equivalence classes of identical instantaneous states what we would normally call gauge orbits are actually sequences of dynamically ordered physically distinct states. The ordering information is specifically encoded in the positivity the multiplier associated with the constraint defined as the lapse above.

In such circumstances the passage to a reduced phase space (where the null directions are quotiented out) will lead to an initial data space without sufficient structure for the reconstruction of dynamics. Here we will give a further argument that this is because in order to construct a solution from an initial data point it essential to also have the ordering information that is abandoned in the reduction procedure (or equivalently through gauge fixing). Formally we can understand this facet of the classical reduction problem in terms of the pull back of the projection to the reduced space not telling us how to parameterise paths in the unreduced space. This point will be further developed by use of the Hamilton-Jacobi formalism for the general case in §9.1.2 and for a free particle model in §9.3.2.

[41]The arguments and mathematics detailed below have entirely joint authorship between myself, S. Gryb. The text of §9.1-9.4 is largely adapted from the published version of the paper (with the material from that paper relating to gravity found within Chapter 15 of this thesis.) The text of §9.5 is based upon my own draft notes the extension of which will from the basis of a further paper. The input of Tim Kowslowski was of great importance in developing these ideas.

In light of the above it is clear that a quantisation procedure predicated upon the reduced phase space correctly parameterising the *full* classical dynamics will be inapplicable for theories with Hamiltonian constraints. This can be show explicitly by consideration a path integral quantisation. See §9.1.1 for the general treatment and §9.3.1 for a pendulum model.

In addition to these arguments stemming from *the problem of reduction* there is good cause to re-evaluate standard quantisation techniques on the grounds of the *the problem of relational time*. As we have seen in §6.1, classical reparameterisation invariant theories can be equipped with an internal and equitable duration measure leading to a fundamentally relational notion of time. Since there are strong conceptual and epistemological arguments in favour of physical models with relational time we would like to be able to construct a quantum theory with relational dynamics. However, as shown in §9.2, standard gauge theory techniques exclude the use of relational clocks and thus temporally relational quantum theories have as yet proved impossible to construct.

In §9.4 we will detail a formal procedure for retaining the essential dynamical ordering information through the introduction of an auxiliary field, and its momenta, that parameterise the classical trajectories and define a relational time. The introduction of these variables is achieved via the extension of the phase space of the original theory in a precise geometrical manner. We then show that the application of standard quantisation techniques to the extended theory will lead to a quantum theory that correctly captures the full dynamics of the original theory we started with. Furthermore, as shown in §9.4.2, in addition to allowing us to retain the full dynamical path integral, the quantisation of the extended theory is such that it leads to a quantum dynamics with respect to a relational time. Our solution is applicable to all theories with global Hamiltonian constraints and thus constitutes a general solution to the global problem of time. §9.5 will explore the implications of relational quantisation for our understanding of observables within reparameterisation invariant theories, and §9.6 will consider some outstanding interpretive issues.

9.1 Diagnosing the problem of time i: the problem of reduction

9.1.1 Gauge invariance versus dynamics: quantum

Globally reparametrization invariant theories feature action functionals in which the integration is performed with respect to an *arbitrary change parameter* λ rather than a fixed Newtonian background time. The invariance of these theories under re–scalings of this parameter leads to the defining feature of their canonical representation: that the Hamiltonian is replaced with a constraint $\mathcal{H}$, often called the *Hamiltonian constraint*. Assuming that all other first class constraints have been gauge–fixed using the method described above,[42] the remaining structure is a phase space $\Gamma(q, p)$ (possibly corresponding to a gauge fixed surface in a larger phase space), a symplectic 2–form, Ω, and a Hamiltonian constraint, $\mathcal{H}$.

We now state the fundamental difference between reparametrization invariant theories and standard gauge theories: the classical solutions are defined as the integral curves of the Hamilton vector field of the constraint $\mathcal{H}$. Because they are the dynamical solutions of the classical theory, the elements of the integral curves of $v_{\mathcal{H}}$ are no more *physically indistinguishable* from each other than this moment is from the big bang. Thus, the leaves of the foliations of the constraint surface, Σ, defined by $\mathcal{H} \approx 0$ can no longer be reasonably identified as gauge orbits - rather they are dynamical solutions. We will now show that, if one turns a blind eye to this fact, one is led to a quantum theory that, in general, cannot contain the appropriate classical limit.

Performing the gauge fixing procedure outlined in §3.3.2, we treat each classical history on the constraint surface as an equivalence class of physically indistinguishable states. We then seek a gauge fixing condition $\rho \approx 0$ satisfying

$$\det |\{\mathcal{H}, \rho\}| \neq 0 \tag{9.1}$$

that selects a single element of each of these foliations. However, because the Hamiltonian is just given by the constraint $\mathcal{H}$, its associated flow is everywhere parallel to the gauge orbits. Thus, this procedure completely trivializes the dynamics since there is no way to flow in any direction on the gauge fixed surface. In addition, the interpretation of the gauge fixed surface is now completely different from the case described in §3.3.2.

[42] Crucially, this is the step that can be performed in shape dynamics that is highly non–trivial in the ADM formulation of general relativity.

Π_{GF} contains a single element of each integral curve of $v_{\mathcal{H}}$. Since each integral curve is itself a possible classical solution, Π_{GF} actually represents a space of initial data for all possible classical evolutions on the constraint surface. Thus, the space, by construction, necessarily excludes any (non-trivial) set of points on the constraint surface corresponding to a classical history. We therefore have that the path integral

$$I = \int \mathcal{D}q_\mu \mathcal{D}p^\mu \delta(\mathcal{H})\delta(\rho) \det |\{\mathcal{H}, \rho\}| \exp\left\{ i \int d\lambda \, [\dot{q}_\mu p^\mu] \right\} \tag{9.2}$$

restricted to Π_{GF} cannot contain any particular solution to the classical evolution problem. It, therefore, *can not* be a quantisation of the original classical theory. This is equivalent to the statement that the Feynman path integral on the reduced space, which is canonically isomorphic to Π_{GF}, fails to capture the classical evolution. In Section 9.3.1, we give an explicit example illustrating this point.

A simple dimensional argument can already be seen indicate this failure to capture the full classical dynamics. In a *typical* gauge theory there are two excess degrees of freedom per first class constraint. Quantisation according to a gauge fixed path integral eliminates this degree of degeneracy since it can be understood as equivalent to the construction of a normal Feynman path integral on the sub-manifold defined by the intersection between the constraints surface and the gauge fixing surface – by definition this sub-manifold has a dimension of two times the number of constraints less than the original phase space. Now, reparameterisation invariance is not a typical gauge symmetry. It corresponds to a freedom in how we label the paths in phase space, results from a symmetry of the action and is also a symmetry of the momenta. It does not, therefore, correspond to a typical gauge symmetry and should not, therefore, be expected to correspond to an excess in the number of degrees of freedom within that space in the usual straightforward manner. Most, importantly, as we have seen, it is no longer correct to think of the phase space as containing null directions that can be treated as unphysical. Formally the only requirement that reparameterisation invariance places upon the phase space is that dynamics be restricted to the constraint surface – and this only equates to reduction by *one* degree of freedom. One conventionally is justified in removing a further degree of freedom (per constraint) only on the grounds of further surplus representational structure existing within the constraint surface: distinct points that are connected by the action of a local symmetry. No such structure is found in the theories in question, so if we press head with the implementation of a procedure that removes *two* degrees of freedom, then we are removing one of them without warrant. Thus, application of a standard gauge fixing quantisation to a reparameterisation invariant theory featuring a single first class Hamiltonian constraint should be expected to lead to a quantum theory in which one of the *physical* degree of freedom is missing. This quantum theory will therefore be unable to reproduce the correct behaviour at the classical limit since it is dimensionally deficient for this purpose.

One key problem we will solve in this chapter will be establishing a consistent quantisation procedure for globally reparametrization invariant theories that *does* contain the appropriate classical limit. We are faced with a dilemma: on one hand, we need to restrict our path integral to a proper symplectic manifold where the Hamilton vector field of $\mathcal{H}$ is well defined on the constraint surface; but, on the other hand, such a restriction must be such that the constraint $\rho \approx 0$ runs parallel to the foliations of $\mathcal{H}$. Unfortunately, this would imply

$$\det |\{\mathcal{H}, \rho\}| = 0 \tag{9.3}$$

and we no longer have a natural candidate for the measure of the path integral. The solution that we will propose in Section 9.4 involves extending the phase space in a trivial way so that the desired classical solutions are indeed contained in the initial value problem of the extended theory. Thus, a standard gauge fixing on this extended theory corresponds to a consistent quantisation of the original theory. Before describing this procedure in detail, we will show how the argument presented above is paralleled in the classical theory.

9.1.2 Gauge invariance versus dynamics: classical

In the semi–classical approximation, the wavefunction of a system is given by the WKB ansatz

$$\psi = e^{iS}, \tag{9.4}$$

where S solves the Hamilton–Jacobi (HJ) equation. When the dynamics is generated by a constraint, the HJ equation takes the form

$$H(q_\mu, \frac{\partial S}{\partial q_\mu}) = 0. \tag{9.5}$$

Hamilton's principal function $S = S(q_\mu, P^a)$ is a function of the configuration variables, q_μ, and the separation constants, P^a. These separation constants are obtained by solving the partial differential equation (9.5). In general, there will be one for each $\frac{\partial S}{\partial q_\mu}$ but these will not all be independent because (9.5) acts as a constraint. This is the reason for labeling P with the index a, which runs from 1 to $d-1$.

The equations of motion are obtained by treating S as a generating function for a canonical transformation from $(q_\mu, p^\mu) \to (Q_a, P^a)$ that trivializes the evolution. The canonical transformation can be determined by computing

$$Q_a = \frac{\partial S(q,P)}{\partial P^a} \tag{9.6}$$

$$p^\mu = \frac{\partial S(q,P)}{\partial q_\mu}. \tag{9.7}$$

S is defined such that the relations (9.7) simply reproduce the Hamiltonian constraint through (9.5). If the system of equations (9.6) can be inverted for q_μ then the equations of motion for q_μ can be determined by using the fact that

$$\dot{Q}_a = 0 \qquad\qquad \dot{P}^a = 0. \tag{9.8}$$

There is, however, an immediate obstruction to this procedure since the system of equations (9.6) has, in general, a one dimensional kernel and, thus, no unique solution. This obstruction can be overcome in two ways:

(i) A *gauge* can be fixed by imposing a gauge fixing condition of the form

$$f(q_\mu, Q_a, P^a, \lambda) = 0. \tag{9.9}$$

f must be chosen such that, when the condition $f = 0$ is imposed, the system of equations (9.6) is invertible.

(ii) The solution space can be parametrized by one of the q's, chosen arbitrarily. This allows us to write

$$q_a = F_a(Q_a, P^a, q_0). \tag{9.10}$$

The first method is the one exclusively employed to conventional gauge theories (i.e., those without Hamiltonian constraints) and is the natural classical analogue of standard methods for dealing with gauge symmetry at a quantum level, in particular the standard Faddeev-Popov gauge fixing methdology of §3.3.2. The gauge fixing (9.9) reduces the dimension of the system. This is natural in standard gauge theory because the map between the original and reduced phase spaces contains no physical information. It is, therefore, reasonable to make the equations of motion invertible by quotienting away the information contained in this map. This does not kill the dynamical information because a non–trivial Hamiltonian survives the quotienting.

For globally reparametrization invariant theories on the other hand, the information contained in the kernel of (9.6) contains *all* the dynamical information. Thus, we must use the second method for reproducing the classical solutions. This is natural, because the relations (9.10) are precisely the integral curves of null directions of the presymplectic form on the constraint surface $H = 0$. In must be noted here that, with regard to this particular point, our analysis is not controversial: method 2 precisely coincides with Rovelli's (Rovelli (2004)) treatment of Newtonian particles (see pp.113-4) and is consistent with how HJ theory is used to reproduce the ADM equations of motion in general relativity – on this see Gerlach (1969). The reader is referred to these sources, and references therein, for more details on this standard treatment. In Section 9.3.2, we will apply methods 1 and 2 to a simple model to illustrate how to implement the formal procedure presented here.

Method 2 is also at least partially related to the correlation scheme discussed in §6.2, since both rely – in subtly different ways – on our ability to isolate one of the variables as providing an internal parameterisation of solutions. Method 2 does not however, entail the interpretation of physical observables as temporally non-local in the non-trivial and conceptually problematic sense of that was discussed in our analysis of the correlation scheme of chapter 6. Moreover, there is no reason to connect the internal parameterisation idea common to both method 2 and the correlation scheme, with the definition of observables as commuting with Hamiltonian constraint – that idea, which, is central to the Rovelli-Dittich-Thiemann notion of observable, should more properly be understood as being implement within the HJ formalism by method 1.

A powerful argument can be made in favour of method 2 over method 1, when dealing with reparametrization invariant theories. In method 1, the pullback under the projection doesn't contain the complete dynamical information. Only in method 2 is it possible to retain information about the temporal ordering of events along the gauge orbits. This a necessary requirement for theories with Hamiltonian constraints because there must be a way to distinguish between the past and the future.[43] This is already implicit in requiring that the lapse, N, should be positive. The fact that *only* Hamiltonian constraints have this

[43]This distinction constitutes a temporal orientation rather than a temporal direction, which would imply an arrow of time.

requirement is an indication that they should be treated differently from the constraints arising in standard gauge theories.

We see that there is a substantive difference between the way the HJ formalism is used in conventional gauge theory and in globally reparametrization invariant theories. This difference is exactly mirrored in the quantum theory. The arguments given in Section 9.1.1 reflect what happens in the classical theory when method 1 is used: the information about the dynamics is lost by quotienting with respect to the null directions of the Hamiltonian flow. A requirement for consistency for the classical and quantum theories is that the method used in the classical theory is reflected in that used in the quantum theory. In light of this requirement and the necessity of using method 2 classically for reparametrization invariant theories, we will present a relational quantisation procedure in Section 9.4.

9.2 Diagnosing the problem of time ii: the problem of relational time

9.2.1 Relational clocks

There is, without doubt, practical utility in the use of a time parameter disconnected from the dynamics of a physical system. Such an external notion of time is an essential element of both Newtonian systems and conventional approaches to quantum theory. Yet, the existence of such a temporal background is inconsistent without the structure of the physical theory that most accurately describes the behaviour of clocks: general relativity. Within this theory, time is an inherently internal notion, parasitic upon the dynamics. Thus, there is empirical motivation to search for a general procedure for consistently constructing an internal notion of time that can be used in both classical and quantum theories.

In addition, there are strong conceptual arguments against external time – many of which predate general relativity. Ernst Mach, in particular, criticized external notions of time on epistemic grounds. In the most general system, we only have access to the internal dynamical degrees of freedom. Thus, it is ‘utterly beyond our power to measure the changes of things by time’ Mach (1960). Rather, according to Mach, any consistent notion of time must be abstracted from change such that the inherently interconnected nature of every possible internal measure of time is accounted for. According to the Mittelstaedt–Barbour (Mittelstaedt (1976); Barbour (1995)) interpretation of Mach, we can understand this *second Mach’s principle* as motivating a relational notion of time that is not merely internal but also equitable; in that it can be derived uniquely from the motions of the entire system taken together. Thus, any isolated system – and, in fact, the universe as a whole – would have its own natural clock emergent from the dynamics. Significantly, for a notion of time to be relational in this sense, it is not enough to be merely internal – it must also be unique and equitable. We cannot, therefore, merely identify an isolated subsystem as our relational clock, since to do so is not only non–unique but would also lead to an inequitable measure, insensitive to the dynamics of the clock system itself.

Within classical non-relativistic theory, relational clocks of exactly the desired type have already been constructed and utilized. As has been pointed out by Barbour, the astronomical measure of *ephemeris time*, based upon the collective motions of the solar system, has precisely the properties discussed above. In Section 9.3, we give an explicit expression for the ephemeris time for a large class of physically relevant finite dimensional models. Quantum mechanically, we run into a problem when attempting to construct a suitably relational notion of time. As we shall discuss in the next section, it is precisely the relational sub-set of internal clocks that are excluded under conventional quantisation techniques. The logic of the next section is as follows: first we establish a general theory for describing evolution in timeless systems in terms of an internal clock as constituted by an isolated subsystem; then, we show that such clocks can never be fully adequate precisely because they are not fully relational.

9.2.2 Internal clocks

We will now detail a method for expressing the path integral (9.2) in terms of evolution with respect to an internal clock constructed from an isolated subsystem. In essence this is an initial attempt to construct the quantum mechanical analogue what we called 'method 2' when discussing the Hamilton-Jaocbi formalism in §9.1.2,. Consider *any* splitting of the Hamiltonian constraint of the form:

$$\mathcal{H} = H_{\parallel} + H_{\perp}. \tag{9.11}$$

Our ability to make this splitting depends principally upon the existence of a sufficiently isolated clock. In practice, the split need only be approximate to some desired order of accuracy. Effectively, we require a clock of the form treated in great detail in Marolf (1995). For more details on the use of internal clocks as a way of modelling relational dynamics and some of the difficulties encountered see Marolf (2009); Giddings *et al.*

(2006); Bojowald *et al.* (2011); Hilgevoord (2005). Given that we have an approximate splitting of the form (9.11), we are in the situation treated in the above references and we can perform a canonical transformation Π

$$\Pi : (q_\mu, p^\mu) \rightarrow (Q_i, P^i, \tau_{\text{int}}, H_{||}) \tag{9.12}$$

generated by the type–2 generating functional $F(q_\mu, P^a, H_{||})$

$$F(q_\mu, P^a, H_{||}) = \int dq_\mu \, p^\mu(q_\mu, P^a, H_{||}). \tag{9.13}$$

The index i runs from $1, .., d-1$. The functions $p^\mu(q_\mu, P^a, H_{||})$ are obtained by inverting the relations

$$\begin{aligned} P^a &= P^a(q_\mu, p^\mu) \\ H_{||}(q_\mu, p^\mu) &= \mathcal{H}(q_\mu, p^\mu) - H_\perp(q_\mu, p^\mu). \end{aligned} \tag{9.14}$$

The functions $P^a(q_\mu, p^\mu)$ are arbitrary provided the above equations are invertible for p^μ. Because $H_{||}$ is fixed by the splitting (9.11), the canonical transformation Π has a $(d-1)$–parameter freedom parametrized by the functions $P^a(q_\mu, p^\mu)$. Up to this freedom, Π singles out an *internal time* variable τ_{int} which can be obtained from

$$\tau_{\text{int}}(q_\mu, p^\mu) = \left. \frac{\partial F}{\partial H_{||}} \right|_{P^a = P^a(q_\mu, p^\mu), H_{||} = H_{||}(q_\mu, p^\mu)}. \tag{9.15}$$

We say that $H_{||}$ singles out an isolated subsystem of the universe whose motion is used as an internal clock parametrizing the motion of the rest of the system. The remaining configuration variables are given by

$$Q_i(q_\mu, p^\mu) = \left. \frac{\partial F}{\partial P^i} \right|_{P^a = P^a(q_\mu, p^\mu), H_{||} = H_{||}(q_\mu, p^\mu)}. \tag{9.16}$$

If we chose to label curves in Γ by the arbitrary parameter λ then, in terms of the transformed coordinates, the natural gauge fixing condition

$$\rho = \tau_{\text{int}} - f(\lambda) \approx 0 \tag{9.17}$$

splits Γ into the two symplectic submanifolds $\Gamma = \Pi_{\text{GF}}(Q_a, P^a) \times \Phi(\tau_{\text{int}}, H_{||})$. From this, it is clear that the freedom in defining τ_{int} through Π corresponds to the unavoidable freedom in making arbitrary canonical transformations on Π_{GF}. Using the simple choice $f(\lambda) = \lambda$, the measure is 1. A short calculation shows that the path integral, I, of (9.2) becomes

$$I = \int \mathcal{D}Q_i \mathcal{D}P^i \exp\left\{ i \int d\tau_{\text{int}} \left[\dot{Q}_a P^a - H_\perp(Q_a, P^a) \right] \right\} \tag{9.18}$$

after integrating over τ_{int} and $H_\parallel$. As expected, it is equivalent to a Feynman path integral on the reduced phase space coordinatized by (Q_a, P^a). This path integral also corresponds to the quantisation of a standard time–dependent Hamiltonian theory in terms of the variables Q^a and momenta P^a with the Hamiltonian $H = H_\perp(Q_a, P^a)$. It leads to a wavefunction satisfying the time–dependent Schrödinger equation

$$i \frac{\partial \Phi}{\partial \tau_{\text{int}}} = \hat{H}_\perp \Psi. \tag{9.19}$$

Furthermore, this is equivalent to applying Dirac quantisation to the Hamiltonian constraint (9.11) after applying Π.[44]

The quantum theory given by (9.18) has a well known classical limit. It is the Hamiltonian theory described by the integral curves of $H_\perp$ parametrized by τ_{int}. Different choices of $f(\lambda)$ correspond to different parametrizations of these integral curves. Although this freedom to reparametrize the classical solutions is a feature we require, the classical solutions we obtain are *not* equal (or equivalent) to the integral curves of $\mathcal{H}$. Instead they are the integral curves of the part projected out of $v_{\mathcal{H}}$ along $\Phi(\tau_{\text{int}}, H_\parallel)$. Thus, the only way to obtain the desired classical limit is to impose $H_\parallel = 0$. *However, with this choice the above method fails since the relations (9.14) become non–invertible*. This is another way of understanding the problem of reduction: gauge fixing such that we follow the integral curves of $v_{\mathcal{H}}$ leads to a zero measure in the path integral. Thus, as it stands, the internal clock methodology is not a true quantum analogue to method 2 of §9.1.2, since it fails to capture the full set of trajectories of the formalism in the semi-classical limit.

In addition to the problem of excluding classical trajectories, this restriction on internal clocks is such that it specifically excludes relational clocks of the kind considered in the previous section. In phase space, the classical ephemeris time is precisely the variable canonically conjugate to the full Hamiltonian of the system. In the following section we demonstrate this explicitly in a simple model.

[44] See Henneaux and Teitelboim (1992) p.280 for a analogous case.

9.3 Toy models

We will now make our formal arguments concrete by applying them to specific models. These models will also help to motivate our new quantisation procedure. We will consider models of the form

$$S = \int d\lambda \sqrt{g^{\mu\nu}(q)\dot{q}_\mu \dot{q}_\nu}, \tag{9.20}$$

where g is some specified metric on configuration space. The variation of this action with respect to q implies that it is a geodesic principle on configuration space. Thus, (9.20) is invariant under $\lambda \to f(\lambda)$, which is the reparametrization invariance we require. These models are useful gravitational models because they include the mini–superspace approximation and contain many key features of general relativity and shape dynamics (See Chapter 15). The fact that these simple models can capture so many features of gravity is often under-appreciated. Indeed, because they correspond to mini-superspace models they are, in fact, genuine symmetry reduced models of quantum gravity. Furthermore, because they are also equivalent to non-relativistic particle models, they have considerable heuristic value.

We will treat the case where g_{ab} is conformally flat. Identifying the conformal factor with $2(E - V(q))$, the Hamiltonian constraint is

$$\mathcal{H} = \frac{\delta_{\mu\nu}}{2} p^\mu p^\nu + V(q) - E \approx 0. \tag{9.21}$$

It is important to note that the origin of this constraint can be traced back to the reparametrization invariance of the action (9.20). As such, its interpretation is crucially different from that of the gauge generating constraints discussed in Part I.

The classical theory corresponding to the Hamiltonian (9.21) is just that of non–relativistic particles under the influence of a potential $V(q)$ with total energy E and mass $m = 1$. The classical equations of motion are easily seen to lead to

$$\sqrt{\frac{E-V}{T}} \frac{d}{d\lambda} \left(\sqrt{\frac{E-V}{T}} \frac{dq_\mu}{d\lambda} \right) = -\frac{\partial V}{\partial q^\mu}, \tag{9.22}$$

where $T = \frac{\delta_{\mu\nu}}{2} p^\mu p^\nu$ is the kinetic energy. If we define the reparametrization invariant quantity

$$\tau_{\text{eph}} \equiv \int d\lambda \sqrt{\frac{T}{E-V}} \tag{9.23}$$

then (9.22) becomes

$$\frac{d^2 q_\mu}{d\tau^2} = -\frac{\partial V}{\partial q^\mu}, \tag{9.24}$$

which are Newton's equations with τ_{eph} playing the role of absolute time. Newton's theory is then given by the integral curves of (9.21) parametrized by the ephemeris time label τ_{eph}.

9.3.1 Example: double pendulum

Consider the double pendulum consisting of 2 particles q_1 and q_2 in 1 dimension under the influence of a potential

$$V(q) = \frac{1}{2}\left(q_1^2 + q_2^2\right) \tag{9.25}$$

corresponding to 2 uncoupled harmonic oscillators whose spring constants k have been set to 1. The Hamiltonian constraint is[45]

$$\mathcal{H}_{\text{HO}} = \frac{1}{2}\left(p_1^2 + p_2^2 + q_1^2 + q_2^2\right) - E. \tag{9.26}$$

Its Hamilton vector field $v_{\mathcal{H}}$ is

$$v_{\mathcal{H}} = p^\mu \frac{\partial}{\partial q_\mu} - q_\mu \frac{\partial}{\partial p^\mu} \tag{9.27}$$

and the constraint surface is the S^3 boundary of the 4–sphere of radius $\sqrt{2E}$. The integral curves on $\mathcal{H} = 0$ are circles when projected into the (q_μ, p^μ)–planes as is familiar from the usual harmonic oscillator.

Performing a standard gauge fixing, as described in §3.3.2, and following the procedure described in Section 9.2.2, we split the Hamiltonian constraint into the pieces

$$H_\parallel = \frac{1}{2}\left(p_1^2 + q_1^2\right) \qquad\qquad H_\perp = \frac{1}{2}\left(p_2^2 + q_2^2\right) - E. \tag{9.28}$$

Using this splitting, we can single out particle 1 as an internal clock for the system. We perform a canonical transformation that takes us to the internal clock variables for particle

[45] In this section we will sometimes write the coordinates of p using lower case indices for convenience.

1 and leaves particle 2 unchanged. The relations (9.14) become

$$p_2 = P_2 \tag{9.29}$$

$$H_{\|} = \frac{1}{2}\left(p_1^2 + q_1^2\right). \tag{9.30}$$

Inverting these, we are led to the generating functional

$$F = \int dq_1 \sqrt{2H_{\|} - q_1^2} + q_2 P_2. \tag{9.31}$$

The transformed Q_2 coordinate is q_2 as expected and the internal time variable canonically conjugate to $H_{\|}$ is

$$\tau_{\text{int}} = \left.\frac{\partial F}{\partial H_{\|}}\right|_{H_{\|}=\frac{1}{2}\left(p_1^2+q_1^2\right)} = \arctan\left(\frac{q_1}{p_1}\right). \tag{9.32}$$

As can be seen from the definitions of τ_{int} and $H_{\|}$ in terms of q_1 and p_1, this canonical transformation takes us to polar coordinates on the (q_1, p_1)–plane of phase space.

The transformed Hamiltonian is

$$\mathcal{H} = H_{\|} + \frac{1}{2}\left(P_2^2 + Q_2^2\right) - E. \tag{9.33}$$

Its Hamilton vector field is

$$v_{\mathcal{H}} = P_2 \frac{\partial}{\partial Q_2} - Q_2 \frac{\partial}{\partial P_2} + \frac{\partial}{\partial \tau_{\text{int}}}. \tag{9.34}$$

The constraint surface is a cylinder along the τ_{int} direction about the (q_2, p_2)–plane of radius $E - H_{\|}$. The integral curves of $v_{\mathcal{H}}$ are helices along the τ_{int}–direction and wrap around the $H_{\|}$–direction implying that $H_{\|}$ is a classical constant of motion (see Figure (9.1)).

If we impose the gauge fixing condition $\tau_{\text{int}} = \lambda$, the path integral (9.18) takes the form

$$I_{\text{HO}} = \int \mathcal{D}Q_2 \mathcal{D}P_2 \exp\left\{ i \int d\tau_{\text{int}} \left[\dot{Q}_2 P_2 - \frac{1}{2}\left(P_2^2 + Q_2^2\right) + E \right] \right\}, \tag{9.35}$$

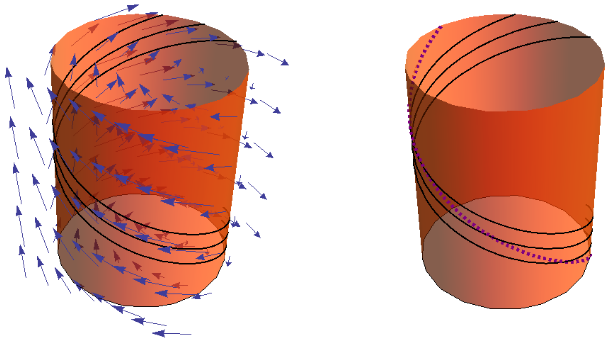


Figure 9.1: The left hand graphic shows the constraint surface $H_{\|} = 0$, the vector field $v_{\mathcal{H}}$ and three examples of classical solutions (integral curves of $v_{\mathcal{H}}$). The right hand graphic shows, as a dashed line, a sample path that is included in the integral (9.35), but, by definition, is nowhere parallel to the classical solutions.

which leads to the time–dependent Schrödinger equation

$$i\frac{\partial \Psi}{\partial \tau_{\text{int}}} = \left(-\frac{1}{2}\frac{\partial^2}{\partial Q_2^2} + \frac{1}{2}Q_2^2 - E\right)\Psi. \tag{9.36}$$

This is the same theory we would have obtained had we quantised the 1D harmonic oscillator in the usual way. However, the freedom in redefining $\tau_{\text{int}} = f(\lambda)$ allows us the freedom to reparametrize the paths in phase space arbitrarily. Although this is the reparametrization freedom we want, it doesn't give the freedom to reparametrize the *full* set of classical solutions.

An easy way to see that these paths will be excluded is to realize that these paths will contribute to the path integral with zero measure because the determinant $\det|\{\mathcal{H}, \tau_{\text{int}}\}|$ is zero for these paths. On the other hand, the paths that *are* captured in the integration are the ones corresponding to the 1D harmonic oscillator when projected down to the (q_2, p_2)–plane. This fact is reflected in our final result. In other words, this gauge fixing has effectively quantised the reparametrization invariant 1D harmonic oscillator, *not* the 2D oscillator we started with.

9.3.2 Example: relational free particle

In this section, we will solve for the classical trajectories of the relational free particle using the HJ formalism. We will compare methods 1 and 2, presented in Section 9.1.2, to show why conventional gauge theory methods should not be used in this case.

For the free particle, the Hamiltonian constraint (9.21) takes the form

$$\mathcal{H} = \frac{\delta_{\mu\nu}}{2} p^\mu p^\nu - E \approx 0. \tag{9.37}$$

Thus, the HJ equation reads

$$\frac{\delta_{\mu\nu}}{2} \frac{\partial S}{\partial q_\mu} \frac{\partial S}{\partial q_\nu} - E = 0. \tag{9.38}$$

This equation can be explicitly solved by introducing the $d - 1$ separation constants P^a. The solution is

$$S(q_\mu, P^a) = q_a P^a \pm \sqrt{2E - P^2} q_0, \tag{9.39}$$

where $P^2 = \delta_{ab} P^a P^b$. We can solve the classical equations of motion by solving for Q_a and p^μ, then by inverting these relations in terms of q_μ. Differentiating S gives

$$Q_a = \frac{\partial S}{\partial P^a} = q_a \mp \frac{\delta_{ab} P^b}{\sqrt{2E - P^2}} q_0 \tag{9.40}$$

$$p^a = \frac{\partial S}{\partial q_a} = P^a \tag{9.41}$$

$$p^0 = \frac{\partial S}{\partial q_0} = \pm\sqrt{2E - P^2}. \tag{9.42}$$

We recover the Hamiltonian constraint, (9.37), immediately from the last two relations for p^μ. As expected, (9.40) is non–invertible for q_μ.

There are two possible ways to deal with the non–invertibility of (9.40):

- **Method 1:** Impose the gauge fixing condition

$$q_0 = \lambda. \tag{9.43}$$

Then,

$$q_a(\lambda) = Q_a \pm \frac{\delta_{ab} P^b}{\sqrt{2E - P^2}} \lambda. \tag{9.44}$$

This does not represent the full set of classical solutions. The reason for this is that, when a gauge fixing is performed, the information about the gauge fixing condition itself is lost. This must be the case, otherwise the theory would not be gauge invariant. Thus, these solutions give curves in the space of q_a's, *not* the space of q_μ's. What is lost is the dynamical information of the gauge fixed variable q_0.

- **Method 2:** We can parametrize the solutions for q_a in terms of q_0, giving

$$q_a(q_0) = Q_a \pm \frac{\delta_{ab} P^b}{\sqrt{2E - P^2}} q_0. \tag{9.45}$$

These are indeed the correct classical solutions as they represent straight lines on configuration space with parameters specified by initial conditions. The two branches of the solution represent the ambiguity of specifying an arrow of time, since our formalism is indifferent to the direction in which time is increasing.

We can straightforwardly see that method 2 is the correct way of reproducing the classical trajectories. However, this method is *not* compatible with standard techniques used for dealing with gauge systems. This is because gauge invariance requires that the gauge fixed theory is ignorant to the details of the gauge fixing itself. This information, however, is necessary for determining the classical trajectories. Thus, it can not be the case that applying standard gauge theory methods to reparametrization invariant theories will lead to the appropriate quantum theory.

9.4 Solving the problem of time: relational quantisation

In the preceding discussion, we have shown how and why standard gauge theory techniques fail to deliver the appropriate quantum theory when applied to theories with global Hamiltonian constraints. According to our diagnosis it is this inappropriateness of the standard canonical quantisation techniques that leads to the problem of time. Our proposed solution to the problem is not to abandon these techniques altogether – to do so would be to deny ourselves access to number of important mathematical results. Rather, we will outline a formal procedure for modifying an arbitrary globally reparametrization invariant theory such that existing gauge methods *can* be applied and lead to the appropriate quantum theory. In doing so, we will also allow for the construction of a class of quantum theories featuring dynamics with respect to a relational time.

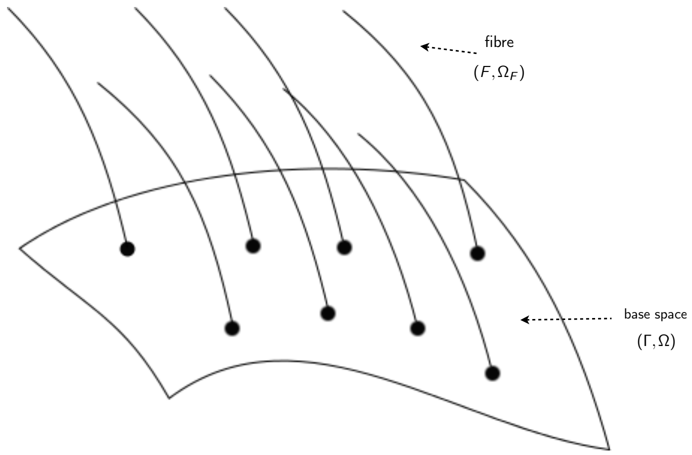


Figure 9.2: This picture shows the fibre bundle structure of the extended theory. The base space is the phase space of the original theory, $\Gamma(q,p)$, and the fibres are two dimensional symplectic manifolds coordinatized by ϵ and τ.

9.4.1 Formal procedure

Consider the general reparametrization invariant theory $\mathcal{T}$ on the phase space, $\Gamma(q,p)$, with symplectic 2–form, $\Omega = dq \wedge dp$, and Hamiltonian constraint, $\mathcal{H}$. We assume that all other first class constraints have been gauge fixed according to the procedure outline in §3.3.2. We define the central element, ϵ, of the Poisson algebra as an observable that commutes with all functions on Γ. As such, ϵ is a constant of motion. Thus, provided we fix its value by observation, it can be added to the Hamiltonian without affecting the dynamics of the theory: $\mathcal{H} \to \mathcal{H} + \epsilon$. In our particle models the central element is simply the total energy of the system and is therefore certainly an observable that can be experimentally fixed. In general relativity, as we will see in Chapter 15, the central element is the cosmological constant.

Now consider the two dimensional symplectic manifold (F, Ω_F) coordinatized by ϵ and its conjugate momentum τ (i.e., $\Omega_F = d\epsilon \wedge d\tau$). We can construct the fibre bundle $(\Gamma_e, \Gamma, \pi_e, F)$ where F is the fibre, Γ is the base space, Γ_e is the fibre bundle itself, and π_e is a continuous surjection $\pi_e : \Gamma_e \to \Gamma$. For our purposes, it will be sufficient to consider a trivial bundle structure so that Γ_e is simply the direct product of Γ with F. The symplectic structure on Γ_e is, thus, given by the non–degenerate symplectic form $\Omega_e = \Omega \times \Omega_F$, which endows Γ_e with a Poisson structure. The overall picture is illustrated in Figure (9.2). We propose that the fibre bundle Γ_e is the phase space, $\Gamma_e(q, p, \tau, \epsilon)$, of an extended theory

$\mathcal{T}_e$ that, when quantised with conventional gauge theory methods, leads to a quantum theory that: **a)** correctly describes the classical solutions of $\mathcal{T}$ in the semi–classical limit without any additional assumptions, and **b)** describes quantum dynamics with respect to a relational notion of time.

We can establish **a)** as follows. First, consider $\mathcal{T}_e = \{\Gamma_e, \Omega_e, \mathcal{H}_e\}$ where $\mathcal{H}_e$ is the extended Hamiltonian constraint

$$\mathcal{H}_e = \mathcal{H}' + \epsilon, \tag{9.46}$$

where $\mathcal{H}'(q, p, \tau, \epsilon)$ is the pullback of $\mathcal{H}(p, q)$ under the bundle projection. Next, consider the constraint surface Σ defined by $H_e = 0$ and the closed degenerate two form $\omega_e \equiv \Omega_e|_\Sigma$. The null direction of ω_e is generated by the Hamilton vector field $v_{\mathcal{H}_e}(\cdot) = \{\cdot, \mathcal{H}_e\}$, where the Poisson structure on Γ_e is used to compute the brackets. This vector field spans the kernel of ω_e and is a codimension 1 submanifold of Σ. Crucially, the kernel of ω_e does not, by definition, contain any physically relevant dynamical information because the null directions have a trivial parametrization given by $v(\tau)$ and are, thus, associated with the physically trivial extension procedure. Thus, a gauge fixing (eg., $\tau =$ const) on Σ corresponds simply to a section on the bundle. Such a gauge fixing selects a gauge fixed surface Π_{GF} that is non–degenerate by construction. It is important to observe that Π_{GF} has the same dimension as the original phase space since the constraint $\mathcal{H}_e$ and the gauge fixing condition each reduce the phase space degrees of freedom by one, thus eliminating the original auxiliary degrees of freedom τ and ϵ. Since the reduced phase space is isomorphic to a section on the bundle, it is also isomorphic to the base space. The classical solutions are contained, therefore, in the reduced phase space because the projection π_e maps

$$\pi_e : v_{\mathcal{H}_e} \to v_{\mathcal{H}}, \tag{9.47}$$

which is the Hamilton vector field of the original Hamiltonian. In Section 9.4.2, we shall first show how relational quantisation is achieved in practice by considering an explicit example and then give a physical interpretation of our result.

9.4.2 Proposed solution

We can establish **b)** as follows. The path integral for this theory is defined using the methods outlined §3.3.2 and using boundary conditions for τ that are consistent with the value of ϵ determined observationally. Note that $\mathcal{H}_e$ is already in the form $\mathcal{H} = H_\parallel + H_\perp$.

Using $H_{\|} = \epsilon$ and $H_{\perp} = \mathcal{H}'$, we can treat τ as an internal clock by imposing the gauge fixing condition $\tau = \lambda$. Thus, τ is an ephemeris clock for the theory $\mathcal{T}$ as it is canonically conjugate to $\mathcal{H}$ under the bundle projection. The gauge fixed path integral is

$$I = \int \mathcal{D}q\mathcal{D}\tau\mathcal{D}p\mathcal{D}\epsilon\delta(\mathcal{H}_{\mathrm{e}})\delta(\tau - \lambda)\det|\{\mathcal{H}_{\mathrm{e}}, \tau - \lambda\}|\exp\left\{i\int d\lambda\,(\dot{q}p + \dot{\tau}\epsilon)\right\} \quad (9.48)$$

Integration over τ and ϵ leads to:

$$I_{\mathcal{T}} = \int \mathcal{D}q\mathcal{D}p\exp\left\{i\int d\tau\left[\frac{dq}{d\tau}p - \mathcal{H}'\right]\right\}, \quad (9.49)$$

which obeys the differential equation

$$i\frac{\partial\Psi}{\partial\tau} = \hat{\mathcal{H}}'\Psi. \quad (9.50)$$

Thus, we obtain a time–dependent Schrödinger equation where the Hamiltonian is the Hamiltonian of the original theory and the time variable τ has a classical analogue corresponding to the total change of the system. We have, therefore, passed to a quantum theory where evolution is with respect to a relational notion of time. Although a convenient gauge choice has been made to write this result there is still freedom to use an arbitrary reparametrization $\tau = f(\lambda)$ as the path integral is invariant under the choice of gauge fixing functions. This implies that the fundamental symmetry of the classical theory is still respected quantum mechanically.

As an example, we can apply this quantisation procedure to the toy model of Section 9.3. The central element ϵ is identified with the negative of the total energy E of the system. We now extend the phase space to include ϵ and its conjugate momentum τ and $E \to -\epsilon$ in the Hamiltonian,

$$\mathcal{H}_{\mathrm{e}} = \frac{\delta_{\mu\nu}}{2}p^{\mu}p^{\nu} + V(q) + \epsilon. \quad (9.51)$$

Using the gauge fixing $\tau = \lambda$, the quantum theory is given by the path integral

$$I = \mathcal{D}q_{\mu}\mathcal{D}p^{\mu}\exp\left\{i\int d\tau\left[\frac{dq_{\mu}}{d\tau}p^{\mu} - \left(\frac{\delta_{\mu\nu}}{2}p^{\mu}p^{\nu} + V(q)\right)\right]\right\}. \quad (9.52)$$

This corresponds to the time–dependent Schrödinger theory

$$i\frac{\partial\Psi}{\partial\tau} = \left[-\frac{\partial^2}{\partial q_\mu^2} + V(q)\right]\Psi = \hat{\mathcal{H}}\Psi. \tag{9.53}$$

In the semi–classical limit, (9.53) reduces to the HJ equation for the phase, S, of the wavefunction

$$\frac{\delta_{\mu\nu}}{2}\frac{\partial S}{\partial q_\mu}\frac{\partial S}{\partial q_\nu} + \frac{\partial S}{\partial \tau} + V(q) = 0. \tag{9.54}$$

We can do a separation ansatz of the form

$$S(q,\tau;P,\epsilon) = W(q,P) - E\tau, \tag{9.55}$$

where $W(q,P)$ solves the equation

$$\frac{\delta_{\mu\nu}}{2}\frac{\partial W}{\partial q_\mu}\frac{\partial W}{\partial q_\nu} + V(q) = E. \tag{9.56}$$

We, thus, recover the usual HJ formalism. This procedure, however, is invariant under $\lambda \to f(\lambda)$ so that we maintain the required reparametrization invariance.

In essence, our proposal is that, for a given reparametrization invariant theory with a single Hamiltonian constraint, we can derive the correct quantum theory by applying the standard quantisation techniques to an extended version of the original theory. It is important to note that this extended theory is merely an *intermediary formalism*: the relational quantum theory that is derived should be understood as constituting the quantum analogue of the original classical theory and not the extended classical theory.

For the simplest class of reparametrization invariant models (including our toy model) – often called *Jacobi's theory* – relational quantisation is equivalent to treating the standard quantisation of a parameterised particle model as the quantum analogue to the classical Jacobi's theory – i.e., the parameterised particle model plays the role of the intermediary formalism.[46] Thus, *mathematically*, the quantum formalism we arrive at is in fact equivalent to that derived (for instance) by Henneaux and Teiteboim Henneaux and Teitelboim (1992) (again see p.280) when considering the quantisation of a parameterised particle theory. However, *physically* our result is importantly different since such authors,

[46]For an elegant treatment of both Jacobi and parameterised particle models the reader is referred to Lanczos (1970)

following the standard approach, consider the quantum analogue of *Jacobi's theory* to be a Wheeler–DeWitt theory without fundamental temporal structure. As shall be explained in Chapter 15, this difference of interpretation with regard to the correct quantisation of simple reparameterisation invariant particle models has important implications within the gravitational context.

Furthermore, in of itself the interpretational shift implied by relational quantisation has subtle but important implications for our understanding of energy within the context of particle models. Whereas, in conventional understandings of Jacobi's theory, energy is interpreted as a constant of nature, and therefore the same for all solutions; within our understanding it becomes a constant of motion that can differ between solutions and is determined experimentally for each solution. This reinterpretation has no classical experimental difference, but rather leads to a quantum formalism that, unlike its conventional rivals, retains the full classical solutions in the appropriate limit.

Our choice is thus between, on the one hand, a conventional quantisation procedure that leads to a quantum formalism with only one energy eigenstate and *does not* allow us to recover Jacobi's theory in the classical limit. And on the other hand, the relational quantisation procedure which leads to a quantum formalism with a classical limit that is *operationally indistinguishable* from Jacobi's theory, but implies a subtly different interpretation of energy at the classical level. Clearly, on this basis one is justified in asserting that it is the relational quantisation option that more faithfully represents a quantum analogue of the classical theory.

It might still be claimed, however, that the formalism we arrive at is merely the quantisation of the extended theory – and our identification of it as corresponding the original theory is not justified. In the following section this point will be dealt with by explicitly showing that the functions treated as observables within the quantisation of the intermediary formalism can be understood as representing the *dynamical* degrees of freedom of the original theory.

9.5 Observables and the intermediary formalism

In this section we will considerer the structure of the observables within the classical reparameterisation invariant theories $\mathcal{T}$ and $\mathcal{T}_e$ above. This treatment will to elucidate the role of the intermediary formalism, in particular with regard to the physical observables which it defines.

Consider a generic reparametrization invariant theory $\mathcal{T} : \{\Gamma(q,p), \Omega, \mathcal{H}(q,p) = 0\}$ and define:

- Real phase space functions: $f \in C^{\infty}(\Gamma) : \Gamma \to \mathbb{R}$
- Poisson bracket structure: $\{f, g\} = \Omega^{ab}\partial_a f \partial_b g$
- Constraint surface: $\Sigma = \{(q,p) \in \Gamma | \mathcal{H}(q,p) = 0\}$
- Weakly vanishing Poisson bracket: $\{f, g\} \approx 0 \leftrightarrow \{f, g\}_{|\mathcal{H}=0} = 0 \leftrightarrow \{f, g\}_W = 0$
- Dirac functions: $d \in C^{\infty}(\Gamma)$ such that $\{d, \mathcal{H}\}_W = 0$

There are three algebras that it is interesting to consider for $\mathcal{T}$:

(i) Poisson algebra $\mathcal{P} : (f, \{,\})$

(ii) Dirac algebra $\mathcal{D} : (d, \{,\}_W)$

(iii) Kuchar-Babour-Foster algebra $\mathcal{KBF} : (f, \{,\}_W)$

Only the third group can be considered a viable candidate to represent the physical observables of $\mathcal{T}$. This is because even though all three groups contain a representation of the Hamiltonian function, the Poission and Dirac algebras render its action either unphysical (in the first case) or trivial (in the second).

Explicitly, within $\mathcal{KBF}_{\mathcal{T}}$ we have that for any f there is an evolution function $\dot{f} = \{f, \mathcal{H}\}_W$ and that $\dot{f} \not\equiv 0$ (i.e., there are always some non-trivial evolution functions). Thus the Kuchar-Barbour-Foster algebra can represent non-trivial dynamics. One should of course also restrict the functions to the constraint surface to ensure only physical states are represented.

Now consider an extended version of our original reparameterisation invariant theory $\mathcal{T}_e : \{\Gamma_e, \Omega_e, \mathcal{H}_e = 0\}$ where we define:

- The central element of $\mathcal{P}_{\mathcal{T}} : \epsilon \in C^{\infty}(\Gamma)$ such that $\{\epsilon, f\} \equiv 0$
- The conjugate variable to $\epsilon : \tau$

- 2D symplectic manifold: (F, Ω_F) with $\Omega_F = d\epsilon \wedge d\tau$
- Extended phase space $\Gamma_e(q, \tau, p, \epsilon)$ as the fibre bundle: $(\Gamma_e, \Gamma, \pi_e, F)$ with the bundle projection being $\pi_e : \Gamma_e \to \Gamma$
- Extended symplectic two form: $\Omega_e = \Omega + \Omega_F$
- Real extended phase space function: $f_e \in C^\infty(\Gamma_e) : \Gamma_e \to \mathbb{R}$
- Extended Poisson bracket: $\{f_e, g_e\}_e = \Omega_e^{ab}\partial_a f_e \partial_b g_e$
- Extended Hamiltonian constraint: $\mathcal{H}_e = \mathcal{H} + \epsilon = 0$
- Extended constraint surface: $\Sigma_e \in \Gamma_e | \mathcal{H}_e(q, \epsilon, p, \tau) = 0\}$
- Weakly vanishing extended Poisson bracket: $\{f_e, g_e\}_e \approx_e \{f_e, g_e\}_e|_{\Sigma_e} = \{f_e, g_e\}_{eW} = 0$
- Extended Dirac functions: $d_e \in C^\infty(\Gamma_e)$ such that $\{d_e, \mathcal{H}_e\}_{eW} = 0$

Again for $\mathcal{T}_e$ there are three algebras of interest:

(i) Poisson algebra $\mathcal{P}_e : (f_e, \{,\}_e)$

(ii) Dirac algebra $\mathcal{D}_e : (d_e, \{,\}_{eW})$

(iii) Kuchar-Barbour-Foster algebra $\mathcal{KBF}_e : (f_e, \{,\}_{eW})$

Our central claim is that the Dirac algebra of the extended theory is equivalent to the physical observables of the unextended theory – these we identify with the relevant $\mathcal{KBF}$ observables. In order to evidence this claim it would seem reasonable that we must establish first that the relevant functions have the relevant relationship in general, and second that there is a robust notion of *evolution* that is preserved between what we are arguing to be representations of the same physical structure. This first issues is fundamentally the question as to whether the Dirac algebra of the extended theory, $\mathcal{D}_e$, is symplectically isomorphic to the Kuchar-Babour-Foster algebra of the unextended theory, $\mathcal{KBF}$. This holds trivially since the bundle projection can be understood as defining the relevant structure persevering map definition.

The key second step is then showing that the bundle projection also preserves weak Poisson bracket structure such that the action of the *unextended* Hamiltonian on *extended*

Dirac obervables projects onto its action on arbitrary unextended functions when evaluated on the original constraint surface:

$$\dot{f} = \{f, \mathcal{H}\}_W = \pi_e(\{d_e, \mathcal{H}\}_{eW}) \tag{9.57}$$

We can show this as follows. First we have

$$\begin{aligned} 0 &= \{d_e, \mathcal{H}_e\}_{eW} && (9.58)\\ &= \{d_e, \mathcal{H}\}_{eW} + \{d_e, \epsilon\}_{eW} && (9.59)\\ \{d_e, \mathcal{H}\}_{eW} &= -\{d_e, \epsilon\}_{eW} && (9.60) \end{aligned}$$

We also have that:

$$\begin{aligned} \{d_e, \epsilon\}_{eW} \approx_e \{d_e, \epsilon\}_e &= \frac{\partial d_e}{\partial q^\mu}\frac{\partial \epsilon}{\partial p^\mu} - \frac{\partial d_e}{\partial p^\mu}\frac{\partial \epsilon}{\partial q^\mu} + \frac{\partial d_e}{\partial \tau^\mu}\frac{\partial \epsilon}{\partial \epsilon^\mu} - \frac{\partial d_e}{\partial \epsilon^\mu}\frac{\partial \epsilon}{\partial \tau^\mu} && (9.61)\\ &= \frac{\partial d_e}{\partial q^\mu}\frac{\partial \epsilon}{\partial p^\mu} - \frac{\partial d_e}{\partial p^\mu}\frac{\partial \epsilon}{\partial q^\mu} && (9.62) \end{aligned}$$

So

$$\begin{aligned} \{d_e, \mathcal{H}\}_{eW} &\approx_e -\frac{\partial d_e}{\partial q^\mu}\frac{\partial \epsilon}{\partial p^\mu} + \frac{\partial d_e}{\partial p^\mu}\frac{\partial \epsilon}{\partial q^\mu} && (9.63)\\ &= \left[-\frac{\partial d_e}{\partial q^\mu}\frac{\partial \epsilon}{\partial p^\mu} + \frac{\partial d_e}{\partial p^\mu}\frac{\partial \epsilon}{\partial q^\mu}\right]_{\mathcal{H}+\epsilon=0} && (9.64)\\ &= \frac{\partial d_e}{\partial q^\mu}\frac{\partial \mathcal{H}}{\partial p^\mu} - \frac{\partial d_e}{\partial p^\mu}\frac{\partial \mathcal{H}}{\partial q^\mu} && (9.65)\\ &= \{d_e, \mathcal{H}\} && (9.66) \end{aligned}$$

Thus, since $\pi_e d_e \approx f$,

$$\begin{aligned} \pi_e(\{d_e, \mathcal{H}\}_{eW}) &= \pi_e(\{d_e, \mathcal{H}\}) && (9.67)\\ &\approx \{f, \mathcal{H}\} && (9.68)\\ &= \{f, \mathcal{H}\}_W && (9.69) \end{aligned}$$

Thus the relevant requirement is met and our interpretation of the classical Dirac observables of the extended theory as representing the physical observables of the unextended theory is justified. Since it is precisely these observables which the relational quantisation procedure treats as fundamental to the quantum formalism (they will be operators on the relevant physical Hilbert space) it is clear that our identification this formalism with the original unextended theory is well justified.

Here we can also note that there is a clear relationship between the definition of classical observables within relational quantisation and the stances with regard to observables in reparameterisation invariant theories discussed in Chapter 6: The classical observables defined in the extension procedure used within relational quantisation are precisely those of the Kuchař-Barbour-Foster approach to observables which we discussed in the context of the emergent time strategy of Chapters 6 and 7.

9.6 Outstanding interpretive issues

Certain key aspects of the implications of the positive results of the last few sections are not completely clear. In particular, it is not entirely apparent what meaning we should assign to either our ephemeris time parameter or our modified TDSE in a quantum cosmological setting. Moreover, given that equitable duration emerges as a purely classical notion – through the correspondence between our ephemeris time parameter and its classical analogue at the semi-classical limit – our analysis may provide insight into the interpretation of classical observations in quantum theory. Furthermore, as in any quantum formalism, there are difficulties in understanding or interpreting both the Born rule probabilities and the relationship of superpositions of states to our phenomenology. Thus, the aspects of the 'problem of time in quantum theory' which are related to the broad category of 'the measurement problem' are still troublesome within our relational quantum formalism.

These interesting and important issues are outside the scope of our current project and will therefore not be further investigated here. So far as non-relativistic theory goes, the above approach does solve the 'quantum problem of time' to the extent that this problem is a problem of how to consistently deal with global Hamiltonian constraints, dynamics, and relational time at a quantum level. The challenge of Part III is to apply what we have learnt to the full theory of general relativity.

Part III

The Relativistic Problem of Time

Guide to Part III

Part III sees our discussion progressing to the relativistic problem of time. Here we will seek to apply the ideas developed within of Parts I and II to full theory of canonical general relativity. At a classical level the major thrust of our analysis will be an examination of the relationship between the fundamental diffeomorphism symmetry of the theory and the ontological status of a linear one-dimensional notion of temporality. It is in this context that we will introduce the constraints of the formalism and within which we shall present three interpretational stances which, in different senses, will amount to denials of time. As in the non-relativistic case an important connection shall be established between a failure of reductionism and the quantum aspect of our problem.

We begin in Chapter 10 with a concise presentation of the canonical formulation of general relativity (§10.1), that is supplemented by an analysis for the relationship with its covariant counterpart (§10.2) and an examination of the role of the Hamiltonian constraints in particular (§10.3). We then proceed to detailing the substance of our first denial of time, first in the context of a motivation taken from standard gauge theory (§11.1) and then in the context of a motivation from a notion of reductive space-time relationalism (§11.2), that builds of the closely related discussion of chapter 7. An argument against the first denial on the basis of dynamical trivialisation will then be presented, together with a rebuttal of the principal line of reasoning that has been employed in its favour (§11.3). Chapter 12 will then present the substance of our second denial on the basis of Machian temporal relationalism (MTR) and the emergent time strategy with which it is assoiacted. After presenting MTR in general terms we will isolate the source of a key problem within its application to canonical general relativity (§12.1). Two possible solutions to this problem will then be evaluated, the first in terms of sophisticated temporal relationalism (§12.2) and the second in terms of a scale invariant formulation of gravity called shape dynamics (§12.3). Chapter 13 will introduce the third denial which is based upon the complete observables scheme that has already been introduced for the non-relativistic case. After a brief restatement of essence of this correlation strategy (§13.1), we proceed to first consider the additional ideas necessary for an application to canonical general relativity (§13.2) and then the philosophical implications with regard to the relative ontological status of space and time (§13.3). Chapter 14 then considers both the implications of the failure of classical reductionism for a Dirac style quantisation of gravity (§14.1), and the prospectus for an alternative route towards quantisation based upon the ideas of the foregoing discussion (§14.2). Chapter 15 consists of some preliminary work towards the application of the ideas of Chapter 9 to the full theory of relativity.

Chapter 10

General relativity and the problem of time

10.1 The canonical theory

Consider the covariant formulation of Einstein's general theory of relativity (in vacuo) according to the Einstein–Hilbert action:

$$S = \frac{1}{\kappa}\int_{\mathcal{M}} d^4x\sqrt{|det(g)|}R = \int_{\mathcal{M}} d^4x\mathcal{L}_{EH} \tag{10.1}$$

where $\mathcal{M}$ is a four-dimensional manifold that we assume to be spatially compact without boundary and to have arbitrary topology, $g_{\mu\nu}$ is a metric tensor field of Lorentzian signature and R is the Ricci scalar. Variation of this action according to the principle of least action leads to the Einstein field equations, the solution of which leads to an expression for the metric tensor. This tensor equips the manifold $\mathcal{M}$ with a geometry, and thus we arrive at the set of Riemannian four-geometries $(\mathcal{M}, g_{\mu\nu})$ that we understand as representing the spacetimes which are nomologically admissible under the theory. As well as providing us with the solutions, the action also gives us a precise methodology for defining the fundamental symmetries of the theory in terms of the Lagrangian $\mathcal{L}_{EH}$ and the Noether symmetry condition:

$$\delta\mathcal{L}_{EH} = \partial_\mu(\epsilon^\mu\mathcal{L}_{EH}) \tag{10.2}$$

which is satisfied for any active variation of the gravitational field variable (i.e., the metric tensor) induced by the infinitesimal coordinate transformations $x^\mu \rightarrow x^\mu - \epsilon^\mu(x)$. The set of all such infinitesimal coordinate transformations forms the group of diffeomorphisms of the manifold $\mathcal{M}$ that we will take to constitute the fundamental local symmetry group of the covariant formalism.[47] Each of these two basic elements to the theory (i.e., solutions and symmetry group) are four dimensional and are understood as corresponding to four-dimensional concepts: spacetimes and symmetries of spacetime. As such, the analysis of either is unlikely to fully elucidate the specific role of time within the theory. Rather, we are better placed to understand the temporal structure of general relativity by passing from the covariant formulation to one that is predicated on space and time rather than spacetime. We shall achieve this by focusing on the canonical formulation of general relativity.

The canonical formulation of general relativity has its origin in the work of Paul Dirac and Peter Bergmann towards towards the construction of a quantum theory of gravity. Important early work can be found in Bergmann (1949) and Dirac (1950), the crucial result was first given in Dirac (1958b) (according to Salisbury (2012) the same Hamiltonian was obtained independently at about the same time by B. DeWitt and also by J. Anderson). Here the formalism will be concisely presented according to the formulation of Arnowitt *et al.* (1960, 1962). We first make the assumption that the manifold $\mathcal{M}$ has a topology which is such that $\mathcal{M} \cong \mathbb{R} \times \sigma$, where σ is a three-dimensional manifold with arbitrary topology that we will again assume to be spatially compact and without boundary.[48] What philosophical significance, if any, we should attach to this non-trivial topological requirement will be discussed in the following subsection. Next we define the *foliation* of $\mathcal{M}$ into hypersurfaces $\Lambda_t := X_t(\sigma)$, where $t \in \mathbb{R}$ and $X_t : \sigma \to \mathcal{M}$ is an *embedding* defined by $X_t(x) := X(t, x)$ for the coordinates x^a on σ. What we are interested in specifically is the foliation of a spacetime, $\mathcal{M}$, into spacelike hypersurfaces, Λ_t – so we must restrict ourselves to arbitrary spacelike embeddings. The lengthly process of decomposing the Einstein–Hilbert action in terms of tensor fields defined on the hypersurfaces and the coefficients used to parameterise the embedding (the lapse and shift below) then leads to a Lagrangian formulation of general relativity in terms of space and time rather than spacetime (see Thiemann (2007) for a full treatment). Finally, recasting this '3+1' Lagrangian

[47] As pointed out by Pons *et al.* (2010), general relativity actually admits the larger symmetry group of field-dependent infinitesimal coordinate transformations, and so $Diff(\mathcal{M})$ is properly a sub-group of the fundamental symmetry group. This difference will not be important for our purposes.

[48] Although fairly standard, this choice of boundary conditions can be seen to have significant impact upon nature of the problem at hand. Under an alternative choice where we assume asymptotically flat spacetimes, the situation with regard to symmetries and the constraints is found to be quite different from that considered below (e.g., Arnowitt *et al.* (1960)). See Lusanna (2011) and future work for more detailed discussion of this point.

formalism into canonical terms gives

$$S = \frac{1}{\kappa}\int_{\mathbb{R}} dt \int_{\sigma} d^3x\{\dot{q}_{ab}P^{ab} - [N^a H_a + |N|H]\} \tag{10.3}$$

Here q_{ab} is a metric tensor field on σ, and P^{ab} its canonical momentum defined by the usual Legendre transformation; N and N^a are arbitrary multipliers called the lapse and shift; H_a and H are constraint functions of the form

$$H_a := -2q_{ac}D_b P^{bc} \tag{10.4}$$

$$H := \frac{s\kappa}{\sqrt{det(q)}}[q_{ac}q_{bd} - \frac{1}{2}q_{ab}q_{cd}]P^{ab}P^{cd} - \sqrt{det(q)}\frac{R}{\kappa} \tag{10.5}$$

with D a covariant derivative with the requisite geometrical properties (see Thiemann (2007)). Here κ is the gravitational coupling constant ($\kappa = 8\pi G$ in units where $c = 1$) and s is metric signature (i.e., $s = -1$ for Lorentzian signature and $s = +1$ for Euclidian signature). These are called the momentum and Hamiltonian constraints respectively and (like typical constraints) can be understood as defining a physical phase space Σ in terms of a sub-manifold (the constraint surface) within the full phase space $\Gamma(q, P)$:

$$\Sigma = \{(q_{ab}, P^{ab}) = x \in \Gamma | H_a(x) = 0; H(x) = 0\} \tag{10.6}$$

Beyond their uncontroversial role in defining this sub-manifold, the interpretation of these constraints is a subtle business. In a *typical* constrained Hamiltonian theory (e.g., see Dirac (1964)), it is assumed that if, as in this case, the constraints are first class (i.e., have a Poisson bracket that vanishes weakly on the constraint surface with all the other constraints) then they should be taken to generate unphysical transformations of the canonical variables and to have their origin directly in the local symmetries of the covariant formalism. The extent to which canonical general relativity is *not* a typical constrained Hamiltonian theory in these senses, as well as the consequent interpretation of the constraints, is the decisive issue that will inform much of our discussion.

10.2 Canonical vs. covariant formalisms

To what extent does the canonical formalism capture the same content as the covariant formulation? We can split this question into two parts: i) is an equivalent set of solutions represented in the space and time formalism as was fixed by the spacetime formalism?; and ii) are an equivalent set of local symmetry transformations implemented on the canonical phase space as were found to hold within the covariant configuration space (i.e., the space of four-metrics)?

Focusing on the first question first. Following Isham (1992) we have the following result: given a Lorentzian spacetime as represented by the geometry $(\mathcal{M}, g)$, then if the constraints (4) and (5) are satisfied on every spacelike hypersurface g will also satisfy the Einstein field equations. Conversely, we can also show that given a $(\mathcal{M}, g)$ that satisfies the Einstein field equations then the constraints will be satisfied on all spacelike hypersurfaces of $\mathcal{M}$. This means that the solutions given by the two formalism are equivalent *provided the covariant spacetime can be expressed in terms of a sequence of space-like hypersurfaces*. This requirement is equivalent to insisting that the spacetimes in question are restricted to be globally hyperbolic (see Geroch (1970)) and is of course directly connected to the topological restriction $\mathcal{M} \cong \mathbb{R} \times \sigma$ that was made in setting up the canonical formalism.

At first sight, this might seem to render the canonical formalism fundamentally inadequate for describing spacetime ontology when compared with the covariant formalism. However, this objection that the requirement of globally hyperbolicity renders the canonical formulation of general relativity representatively deficient in comparison with the covariant formalism should not be overstated. By insisting that our spacetime is globally hyperbolic we are only requiring the existence of a Cauchy surface in $\mathcal{M}$, meaning that the only solutions that have been excluded are those inconsistent with the basic notions of causality and determinism that we would *prima facie* have expected to hold within a classical theory anyway. The physical content of non-globally hyperbolic solutions seems hard to countenance since they include strange objects such as closed time-like curves, and in terms of the well confirmed *empirical* content of the theory nothing has been lost since all observational data from currently observed regions of the universe is consistent with the exclusive existence of globally hyperbolic solutions.[49]

Within quantum gravity there is the possibility that one may need access to different kinds of topologies – or, in fact, perhaps even topology changes! However, the reasoning behind lifting the topological requirement at a quantum level does not impinge on its classical status. To the extent to which the solutions of covariant general relativity can be

[49]For an analysis of the connection between non-globally hyperbolic spacetimes, closed time-like curves and the possibility of time machines see Smeenk and Wuthrich (2009).

understood as representing 'physically reasonable' spacetimes, the solutions of canonical general relativity can equally be thought of representing these spacetimes (albeit in terms of space and time).

Still, it *is* true that canonical general relativity is well defined on a set of solutions that is a subset of those of covariant general relativity. This means that in moving from one formalism to the other we are removing from our theoretical toolkit the ability to represent a class of *nomologically possible* worlds. Furthermore, although these worlds might *seem* unreasonable because of their strange causal structure, to exclude them in principle from a philosophical analysis of the nature of time in general relativity would be seriously begging the question – we cannot merely appeal to a view on the nature of time to fix our view on the nature of time! Asserting a causal censorship condition that excludes the non-globally hyperbolic solutions as an additional law of nature is a highly non-trivial move which we will not here propose to make. Rather, one straightforward option is to invert the supposed deficiency into a strength and make global hyperbolicity a prediction rather than a restriction. Such a move depends on our ability to see the canonical formalism on an equal footing to its covariant counterpart, and not as purely parasitic upon it. This we can do by pointing to the fact that – as shown by Hojman *et al.* (1976) – it is possible to derive canonical general relativity without passing through the covariant formalism. Such a manoeuvre will be crucial to the assessment of our second denial of time and will be further discussed in §12.3. Alternatively, we can simply fall back on a weakened stance: this analysis and its conclusions with regard to the nature of time carry with them a global parenthesis of *given the restriction to the globally hyperbolic subset of solutions*. Since our principal object is to examine issues relating to diffeomorphism symmetry and the ontological status of a linear one-dimensional notion of temporality, our discussion can bear this qualification without any undue burden or inconsistency.

More significant to our purpose is the relationship between the respective local symmetry transformations of the two formalisms. Whereas, as discussed above, the covariant action is invariant under the full set of spacetime diffeomorphisms $Diff(\mathcal{M})$, in the canonical formulation it is only a subset of these that is realised. This subset can be shown (e.g., Pons *et al.* (2010)) to be infinitesimal coordinate transformations $x^\mu \to x^\mu - \epsilon^\mu(x)$ such that

$$\epsilon^\mu(x) = n^\mu(x)\xi^0 + \delta^\mu_a \xi^a \tag{10.7}$$

where $n^\mu = (N^{-1}, -N^{-1}N^a)$, and here the ξ^μ are taken to be arbitrary functions of the coordinates. From the perspective of the *derivation* of canonical from covariant general relativity, the origin of this discrepancy between the symmetry transformations realised in the two formalisms is well understood – it can be explained in terms of the spacelike nature of the otherwise arbitrary embedding (see Isham and Kuchař (1985b)) or (relatedly) in terms of incomplete *projectability* between the symmetry transformations defined in the relevant tangent and cotangent bundle structures (see Pons *et al.* (1997)).

Alternatively, we can consider the elegant and important derivation of these canonical symmetry transformations purely in terms of a *deformation algebra* pertaining to spacelike hypersurfaces embedded in a Riemannian spacetime (Teitelboim (1973); Hojman *et al.* (1976)). This treatment gives us a basis for the canonical symmetry transformations independent of the covariant theory and implies that we can understand them as encoding physical content not purely parasitic upon four-dimensional diffeomorphism symmetry. Crucially, this treatment also makes clear the deep connection between the form of the constraints and the nature of the symmetries. In fact, what is shown is that this canonical symmetry group (known as the Bergmann–Komar group $\mathcal{BK}$) is and must be generated by constraints of the specific form (10.4-5), which will inevitably satisfy the constraint algebra

$$\{\vec{H}(\vec{N}),\vec{H}(\vec{N'})\} = -\kappa\vec{H}(\mathfrak{L}_{N_a}N'_a) \tag{10.8}$$

$$\{\vec{H}(\vec{N}),H(N)\} = -\kappa H(\mathfrak{L}_{N_a}N) \tag{10.9}$$

$$\{H(N),H(N')\} = s\kappa\vec{H}(F(N,N',q)) \tag{10.10}$$

where $H(N)$ and $\vec{H}(\vec{N})$ are smeared versions of the constraints (e.g., $\vec{H}(\vec{N}) := \int_\sigma d^3xN^aH_a$) and $F(N,N',q) = q^{ab}(NN'_{,b} - N'N_{,b})$. The presence of structure functions on the right-hand side of (10.10) means that strictly $\mathcal{BK}$ is not a group (and the constraint *algebra* is not technically an algebra) and of course emphasises that $Diff(\mathcal{M}) \neq \mathcal{BK}$.

Despite these important differences, the symmetries of the covariant and canonical formalisms can in fact be shown to be physically equivalent since given a solution to the equations of motion within the canonical formalism the action of $\mathcal{BK}$ will coincide with that of $Diff(\mathcal{M})$ (for diffeomorphisms connected with the identity). Thus, at the classical level at least there is no detectable difference. We can in fact explicitly construct a canonical symmetry generator acting on the space of solutions that maps between the same diffeomorphically related spacetimes that we take to be symmetry related in the covariant formalism (see Pons *et al.* (2010) and references therein for more details). Thus, so long as we are concerned with globally hyperbolic solutions, the two formalisms admit identical symmetry relations.

10.3 Time and the riddle of the Hamiltonian constraints

In the discussion above, the constraints of the canonical formalism were found to be involved in fixing both the dynamics and symmetries in accordance with the physics of covariant general relativity. This dual symmetry/dynamics aspect leads to much confusion and complexity with regard to the constraints – in particular the Hamiltonian constraint – and is at the root of the problem of time in canonical gravity. Whereas the momentum constraints can be understood unambiguously as implementing infinitesimal three-diffeomorphisms *on phase space*, the role of the Hamiltonian constraints in this context is far more opaque. We can see this explicitly by considering the form of the Poisson brackets between each constraint and the canonical variables. For the momentum constraints it takes the form

$$\{\vec{H}(\vec{N}), q_{ab}\} = \kappa(\mathfrak{L}_{\vec{N}} q_{ab}) \tag{10.11}$$

$$\{\vec{H}(\vec{N}), P^{ab}\} = \kappa(\mathfrak{L}_{\vec{N}} P^{ab}) \tag{10.12}$$

The appearance of the Lie derivative on the right-hand side of each equation indicates that these constraints can be understood as generating purely infinitesimal diffeomorphisms of the phase space variables. In fact, it means that, *on their own*, these constraints can be understood as implementing the Lie group of diffeomorphisms of σ (Isham and Kuchař (1985b)).

The Hamiltonian constraints in, stark contrast, have a phase space action that seems, *prima facie*, manifestly *dynamical*. For any specification of the lapse, they effect an infinitesimal phase space transformation from the canonical variables that characterise a given three-geometry to those describing a second three-geometry which is dynamically subsequent. More careful analysis however reveals a dual role within which the seeds of our conceptual enigma are sown. We can consider the explicit action of the Hamiltonian constraints on an *embedded* canonical momentum variable. Such a variable is so called because it is the canonical conjugate of a metric variable $q^{\mu\nu}$, which is a tensor field (the first fundamental form) defined on the embedded hypersurface Λ_t. This new metric variable can be expressed purely in terms of spatial vector fields on Λ_t and the usual metric variable on σ, q^{ab} (see Thiemann (2007, Eq. 1.1.16)). The new momentum variable can be written in terms of $q^{\mu\nu}$ together with it and another spatial tensor field on Λ_t (the second fundamental form). An elegant calculation by Thiemann (2007, pp.54-6) yields the explicit expression

$$\{H(N), P^{\mu\nu}\} \;=\; \frac{q^{\mu\nu}NH}{2} - N\sqrt{\|q\|}[q^{\mu\rho}q^{\nu\sigma} - q^{\mu\nu}q^{\rho\sigma}]R^4_{\rho\sigma} + \mathfrak{L}_{Nn}P^{\mu\nu} \quad (10.13)$$

with $R^4_{\mu\nu}$ the Ricci 4-tensor. The first term on the right hand side is zero on the physical phase space (defined by satisfaction of the constraints) and is therefore unimportant. The second is zero for solutions to the equations of motion, and thus we have that *on shell* the Hamiltonian performs the role of generating infinitesimal diffeomorphisms. Whereas the diffeomorphisms associated with the momentum constraints can be understood as purely kinematical symmetries of the three geometries σ (irrespective of whether the equations of motion hold), those associated with the Hamiltonian constraints are symmetries not only of entire spacetimes, but of spacetimes that are solutions.

For a given solution and an embedded hypersurface, the constraints generate a local deformation of the hypersurface. Collectively such an action is equivalent to the refoliation of spacetime and therefore to the generation of a different unphysical *splitting* of the spacetime into space and time. However, the solutions themselves are consequences of the dynamical role that the Hamiltonian constraints play in the context of three-geometries considered on their own rather than as embedded in a spacetime. Thus, to maintain both the fundamental symmetry of the theory *and* the dynamics we must appreciate the dual, context dependent role of the Hamiltonian constraints.[50] In the remaining discussion, we will explore the narrow path that traverses the folly of failing to appreciate either side of this duality and, after observing the perils of falling into the abyss below, we will come upon a fork that forces us to choose between retaining a weaker *Machian* notion of time at the cost of global scale and dispensing with time altogether.

[50] It is important to note that this key aspect to our analysis represents a departure from *both* the received and dissenting view on this matter (although it is close to the spirit of Pons *et al.* (2010)). Whereas, the received view is that the Hamiltonian constraints purely generate unphysical transformations (e.g., Rovelli (2004)), the dissenting view (which could be understood as being defended in Barbour and Foster (2008) and also associated with Kuchař (1991b, 1992)) is that the constraints' action is purely physical in character. Barbour, however, 'suspects that the action of the Hamiltonian constraints in GR is part physical and part gauge' (Julian Barbour, personal communication 2012).

Chapter 11

Denial I: reductive temporal relationalism

11.1 Gauge theory and indeterminism

Motivation for a particularly influential (but ultimately unpersuasive) argument towards the denial of time in canonical general relativity derives from the consideration of the otiose representative structure constituted by the Hamiltonian formulation of a standard gauge theory of the type considered in Part I. We will briefly repeat some of the material already given §1.5 to refresh the readers mind of the relevant general argument, without making any particular reference to the Hamiltonian constraints of canonical general relativity and the important subtleties that go along with them.

Consider a constrained Hamiltonian theory constituted by a phase space Γ parameterised by n canonical coordinates (p, q), a Hamiltonian functional $H(p, q)$ and a set of m first class constraints $\phi_i(p, q) = 0$. Such a theory corresponds to a physical system with gauge freedom if and only if the action is invariant under some local symmetry group. Geometrically we can characterise such a generic constrained Hamiltonian theory in term of a phase space Γ with a symplectic geometry (Γ, Ω) and an $n - m$ dimensional submanifold $\Sigma = \{(p, q) \in \Gamma | \forall_i : \phi_i(p, q) = 0\}$ with a presymplectic geometry (Σ, ω) called the constraint surface or physical phase space. The degenerate structure of the latter can be understood in terms the integral curves of the vector fields that make up the null vector space (or *kernel*) of ω partitioning Σ into a set of transverse sub-manifolds called *gauge orbits*. In physical terms, each of these orbits are assumed to have the significant feature that all of its constituent points are physically indistinguishable – they correspond to an identical value of the Hamiltonian functional (accompanied by equal value for all other measurable quantaties). Furthermore, paths in the physical phase space that differ by a transformation along a gauge orbit will necessarily correspond to identical values of the canonical action and will therefore also be indistinguishable.

Given that we make the usual interpretation of points in the (physical) phase space as representing distinct instantaneous physical states, the above feature is a form of indeterminism (or underdetermination) since given an initial specification of physical states the formalism does not fix a unique continuation. This would seem unsatisfactory for the case of classical theories, including general relativity, where the relevant measurable quantities are manifestly deterministic. The natural response to such circumstances is to point to inadequacy within our representative formalism rather than the characterisation of the connection between what is real and what is measurable within our theories. We assert that there is 'surplus structure' within our formalism as embodied precisely by the directions defined by the gauge orbits. The most obvious methodology for controlling this excess is to classify these directions as unphysical and use points in the space of gauge orbits to give us a unique representation of physical states.[51] Formally we may construct this space of gauge orbits or reduced phase space in terms of the quotient manifold that results from the application of a *symplectic reduction* procedure to the physical phase space Σ. For simple constrained Hamiltonian theories, this reduction is effected simply by taking the quotient Σ by the kernel of ω – see Gotay *et al.* (1978) for the more complex case. Either way, it can be proved for a large class of cases (Souriau (1997)) that the space of gauge orbits that results from the application of symplectic reduction to the physical phase space has a symplectic geometry (Π_R, Ω_R) and inherits a Hamiltonian functional from the physical phase space. Since we have removed the null directions Π_R has a non-degenerate structure and is not, therefore, afflicted with the kind of indeterminism mentioned above. If we assign to points in the reduced phase space the role of representing unique physical states, then the formalism is now such that any initial specification will also imply the provision of a unique continuation.

By passing to the reduced phase space of a constrained Hamiltonian theory, we reap the reward of a formalism trimmed of any superfluous representative structure. This has led some authors to argue that we should endow the reduced space with a privileged status. In particular Gordon Belot and John Earman (Belot (2000, 2003); Earman (2003); Belot and Earman (1999, 2001)) have argued that we should consider the reduced phase space as the fundamental dynamical arena of a gauge theory. As applied to a generic gauge theory, this form of *reductionism*, although open to a number of philosophical objections (see

[51] Less obviously we might instead weaken the representative relationship between points and states via the introduction of some notion of anti-haecceitism. This strategy will be examined carefully within the particular context of time in canonical general relativity as discussed in §12.2

Chapter 2), is a viable option and is to a large degree supported by the various techniques of canonical quantisation for gauge theories – all of which might be seen to be predicated on the reduced phase space (see Chapters 3 and 4). However, for the specific case of the Hamiltonian constraints that feature in canonical general relativity (and as we have seen for non-relativistic reparameterisation invariant theory) such reductionism rests on an inappropriate interpretation of the Hamiltonian constraints as pure gauge generators. We will examine this crucial issue more carefully after first giving a second motivation for symplectic reduction that is specific to canonical general relativity and is based on a form of relationalism appropriate to the spacetime concepts found within the theory.

11.2 Reductive spacetime relationalism

Again reiterating key material from above (in particular Chapter 7) for the convenience of the reader. The philosophical doctrine of relationalism with regard to space and time has its roots in the early modern natural philosophy of Descartes, Leibniz and Huygens but (arguably) takes its most precise form in the work of Mach and Poincaré (Barbour (2009)). In essence, it is a position as to the relative ontological status of relations between material bodies and the entities or objects constituted by space and time themselves. A relationalist is taken to hold that the relations are primary and that space and time are merely derived or abstracted from them. In the context of a theory containing a concept of dynamical spacetime such as general relativity, it is not entirely clear what relationalism as it was originally conceived should be taken to mean, and the modern philosophical discussion is replete with positions that are taken to be either pro- or anti- some version of *relationalism*. Our purpose here will not be to survey this literature or explicitly analyse its connection with the indeterminism issue of the previous section in terms of the famous hole argument.[52] Rather, we will concern ourselves with the notion of reductive spacetime relationalism that is presented by the relevant authors in their argument towards our first denial of time. A second, importantly different notion of temporal relationalism will be discussed in Chapter 12.

Among others, Belot and Earman (1999, 2001) hold that the essence of spacetime relationalism within general relativity should be taken to be the denial of a fundamen-

[52] In essence the hole argument relates to the indeterminism born of the four-dimensional diffeomorphism invariance and is constructed within the covariant formulation of general relativity. See Norton (2011), Rickles (2008) and references therein.

tal ontological role for spacetime *points*. Such points are of course represented within a covariant formalism by the coordinatisation of the manifold $\mathcal{M}$ and will therefore be given distinct representations within different coordinatisations. If we assume that cross identification between points within qualitatively identical spacetime models – i.e., with the same geometry – can be taken to ground a real difference between these models (i.e., they may differ solely haecceitistically), then relationalism can be understood in terms of the denial of exactly such difference on the grounds that spacetime points do not have a fundamental ontological status.[53] A spacetime relationalist is thus someone who will 'deny that there could be two possible worlds with the same geometry that differ only in virtue of the way that is geometry is shared out over existent spacetime points' (Belot and Earman (2001, p.18)). In the context of the covariant formalism, this means that two geometries $(M, g_{\mu\nu})$ and $(M', g'_{\mu\nu})$ that solve the Einstein field equations and are related by an element of $Diff(\mathcal{M})$ are considered to be the same physically possible situation. This is because the difference between them is exactly in terms of the coordinatisations rather than the geometrical structure; therefore the ontologies which they are taken to represent can differ (if they differ at all) only with respect to the role played by the spacetime points. By endorsing such a *Leibniz equivalence* type principle, Belot and Earman disavow this difference.[54] Furthermore, by cutting down the class of distinct possibilities to include only geometries that are members of different diffeomorphically related equivalence classes we have implicitly performed a reduction with respect to our fundamental representative space. Rather than considering the space of Riemannian four-geometries corresponding to four-metrics that solve the Einstein field equations as our basic arena for representing the world, we instead should consider the quotient of that space by the group of four-dimensional diffeomorphisms. Thus, we can see Belot and Earman's arguments as leading us from relationalism to reduction: they are *reductive relationalists*.

What does this reductive form of spacetime relationalism mean in the context of the canonical formalism? If we focus our attention on the role of *spatial* points, then we have a clear answer. In analogy to the spactime case, spatial points are represented in terms of the coordinatisation of a manifold, in this case the three-dimensional manifold σ. Furthermore, the action of the theory is invariant under the group of three-dimensional diffeomorphisms of this manifold, $Diff(\sigma)$, and so a reconstruction of the argument above

[53] We will, for the time being, defer the discussion of histories anti-haecceitist variant of relationalism. See §12.2.

[54] There are, of course, other ways of formulating such a principle that do not have the same implications for 'possibility reduction'. For instance, that suggested by Saunders (2003)

can be made for this case. Explicitly, since two canonical solutions that differ solely on the basis of the application of an element of $Diff(\sigma)$ are physically identical, asserting the existence of spatial points will violate a Leibnizian type principle of equivalence of exactly the same type as that introduced via the quote from Belot and Earman (2001) above. Thus, a reductive relationalist with regard to space will endorse a reduced space of three-geometries as representatively fundamental within the canonical formalism. Since $Diff(\sigma)$ is implemented on phase space by the action of the momentum constraints (see discussion surrounding equations 10.11 and 10.12 within §10.3), we know that precisely the reduced space we are looking for can be achieved by quotienting out the gauge orbits associated with those constraints according to a symplectic reduction procedure above. This, in fact, leads us directly to the phase space equivalent of Wheeler's superspace (see Wheeler (1968); Giulini (2009)) – 'super-phase-space', on which a formulation of canonical general relativity would be constituted according to this brand of spatial reductive relationalism.

With regard to time things are, as ever, far more complicated. One might hope to translate a position of spacetime relationalism as expressed in terms of the covariant formalism into a position of spatial relationalism plus temporal relationalism as expressed in terms of the canonical formalism. Building on the ideas of the previous paragraph, we would hope to disavow the fundamental status of temporal points by enforcing ontological equivalence between solutions that differ only as to the way in which the four-dimensional geometrical structure is 'shared out' over these points. Thus, we would in effect perform a reduction of paths in super-phase-space such that those that differ only as to how time is labelled are classified as the same path; we would arrive at a new *doubly reduced* representative space. Unfortunately, such a naive implementation of reductive temporal relationalism is neither possible nor adequate to our purpose.

Although we have assumed that the spacetime manifold $\mathcal{M}$ has a topology which is such that $\mathcal{M} \cong \mathbb{R} \times \sigma$ and therefore that the temporal dimension is represented in terms of the real line, the complication of foliation invariance means that the arbitrariness with regard to time is not fully captured merely by global temporal relabelling – i.e., by the one-dimensional diffeomorphism group $Diff(\mathbb{R})$. Furthermore, unlike spatial diffeomorphisms, these 'temporal diffeomorphisms' have no representation at the level of constraints acting on phase space points or for that matter even phase space paths, and so it is impossible (in the conventional formalism) to frame this naive temporal relationalism simply in terms of a reduction procedure.

Foliation invariance means that the theory is invariant under the set of *local* temporal relabellings of each point on each space-like hypersurface. The global temporal relabellings discussed above form only a subset of these. To be consistent with the notion of reductive spacetime relationalism defined above, it is the temporal points that constitute these local labellings that must be excluded from our ontology via a Leibniz equivalence inspired quotienting operation. In the case of local temporal relabelling (unlike the global case), we do have a canonical constraints that can be associated with the relevant symmetry: the Hamiltonian constraints. However, as discussed in §10.3 above, the connection between these constraints and refoliation symmetries can only be made precise in the context of spacetimes (corresponding to paths in the physical phase space) that are also solutions to the equations of motion. In the context of their action on *phase space*, the Hamiltonian constraints generate evolution. Thus, although it might seem at first sight that reducing out the action of the Hamiltonian constraints (on the cotangent bundle over superspace – i.e., super-phase-space) will achieve the object of reductive spacetime relationalism within the canonical formalism, our understanding of the constraints' dual role leads to immediate scepticism on this count. The object of reductive relationalism with regard to time is to construct a representative arena in which the distinct possibilities entailed by the existence of temporal points have been removed. Within the context of canonical general relativity, such a reduction makes sense (at least in principle) at the level of entire histories related by a refoliation symmetry. However, it is difficult to see how it can possibly be achieved by a reduction of phase space since such symmetries cannot be represented in terms of the relationship between points on this space. It is exactly this kind of phase space reduction with regard to the Hamiltonian constraints that Belot and Earman argue implements their reductive spacetime relationalism within the canonical formalism and to which we now turn.

11.3 Dynamical trivialisation and the isomorphism argument

We thus have two distinct but connected motivations for enacting a symplectic reduction of the phase space of canonical general relativity with regard to the Hamiltonian constraints. Firstly, we have the argument from indeterminism and surplus structure – it is assumed that as for the case of other theories with first class constraints, the sub-manifolds defined by the integral curves of the null vector fields associated with the Hamiltonian constraints will form gauge equivalence classes. Thus, as for the generic case, the unreduced formalism will possess an excess representation of physical states such that an initial specification of phase space points will admit multiple physically identical but mathematically distinct continuations. By reducing out the action of these constraints we will remove both this indeterminism and the redundant representative structure that enables it. Secondly, we have the motivation from reductive spacetime relationalism – we wish to reduce our possibility space such that differences entailed by distinct coordinatisations of the same fundamental geometrical structure are no longer encoded. Specifically, in addition to removing the representation of spatial points, via reduction with respect to the momentum constraints, we also want to remove the local temporal labellings that play the role of representing temporal points as basic structures within the theory (this second motivation is of course directly connected to the reductive temporal relationalism discussed in the context of non-relativistic theory within Chapter 7).

It is for both these reasons that Belot and Earman (2001, 17-18) advocate the use of a literal reading of the reduced phase space of general relativity – it is the use of this space that they claim allows us to both 'avoid indeterminism' and, in doing so, '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'. They explicitly make the claim that, modulo positions built upon anti-haecceitism (e.g., sophisticated substantivalism. See §7.1), 'one must be a [reductive] relationalist in order to give a deterministic interpretation of general relativity'. Thus, Belot and Earman's reductive spacetime relationalism is directly connected to the interpretation of the reduced phase space as the fundamental dynamical arena. In a canonical context it is *claimed* to be a position such that only the distinct possibilities entailed by the existence of spacetime points have been removed. If 'points of the reduced phase space are just the equivalence classes of diffeomorphic models of general relativity' then, *prima facie*, the only temporal structure we have removed in passage to the reduced space should be the temporal points – and this would not make for a particularly strong denial of time. However, closer analysis reveals that the reduction has in fact removed far more temporal structure from our formalism and thus that the reductive spacetime relationalism of Belot and Earman inevitably leads to a far stronger denial of time.

Like in other gauge theories, the construction of a reduced phase space involves quotienting out of the action of the first class constraints and thus, for the gravitational case, would involving treating the phase space action of $H(N)$ as purely symmetry generating. However, as detailed in §10.3 and mentioned above the role of the Hamiltonian constraints within canonical general relativity is essentially a dual one. When considered as acting on purely on a three geometry (as represented by a phase space point) they generate dynamical evolution and when considered as acting on space-like hypersurface embedded within a solution they generates infinitesimal diffeomorphisms. The latter role means that the constraint can be considered responsible for generating refoliation symmetries and allows us to understand how the four dimensional diffeomorphism symmetry is (to a certain extent) implemented canonically. However, the former role cannot be discounted since without it the solutions within which the hypersurfaces are embedded cannot be defined. Moreover, the *gauge orbits* associated with the constraints action on phase space are in fact closer in character to solutions themselves and are explicitly *not* equivalence classes of solutions since a point with the orbit is associated with a three not four dimensional object. Still, by (erroneously) classifying all phase space points within these sub-manifolds as representing the same *state* we will ensure that any pair of three geometries which are contained within solutions related by a refoliation symmetry will be (again, erroneously) classified as equivalent. Thus symplectic reduction will remove the indeterminism related to that symmetry. It will, of course, therefore additionally mean that the reductive temporal relationalist desire to pass to a representative space that excludes distinct local temporal labellings will also have been achieved.

In addition to these two primary goals, however, this reduction has the dire unintended consequence that all dynamically related three geometries are classified as representing the same state. This is because the orbit that is quotiented is, as it must be the nature of the Hamiltonian constraints phase space action, composed of every state that can be accessed via the 'many fingered' time evolution the theory allows for in terms of the action of the Hamiltonian and the arbitrariness of the lapse function. By reducing the representative capacity of the orbit down to a single state we pass from many fingers to no fingers – and not one finger! Furthermore, since we have not respected the dynamical role of the Hamiltonian, in a phase space context by passing to the reduced space we will have classified states which are *physically distinct* members of a given solution as identical. This *is exactly* to treat the current state of the universe and its state just after the big bang as identical (contra the claims of Belot (2007, p.78)).

We can establish that the reduced phase space of general relativity has the described structure explicitly by considering the following argument based upon an extension of the geometric formalism introduced in §1.5 and reiterated above (for more details of the formalism in question see (Rovelli, 2004, §3.3.2 and §4.3):

- I) The physical phase space of canonical GR will have a presymplectic geometry (Σ, ω) and like in a *typical* gauge theory this structure contains characteristic 'null directions' the 'integral curves' of which we usually identify as gauge orbits.

- II) Since general relativity is a field theory the geometric structure we are dealing with is a little more complex than that introduce before: motions are now four-dimensional surfaces with a quadritangent X (made up of the tensor product of four independent tangents) defined at each point. Thus, the part of 'null direction' is now played by the a specific quadritangent and that of 'integral curve' by an integral surface. ω is still a presympletic form but now it is a five, rather than two form; derived as it is from a canonical four, rather than one form.

- III) More significantly, and in stark contrast to the standard gauge theory case, if we define the gauge orbits of canonical general relativity explicitly in terms of the four dimensional surfaces $\bar{\gamma}$ in Σ such that the quadritangent to the orbit X is in the kernel of ω (i.e., $\omega(X) = 0$), then we can identify the $\bar{\gamma}$ with the set of (globally hyperbolic) solutions of the Einstein field equations (Rovelli, 2004, p.157).

- IV) Since these orbits are precisely those which we would normally classify as gauge equivalence classes a symplectic reduction procedure would (in principle) lead to a reduced phase space within which, *prima facie*, dynamics has been gauged out.

- V) Furthermore, since this reduced space is only equipped with a trivial Hamiltonian function there is no hope of recovering dynamical evolution in terms of transformations between points in the reduced space.

Reductive relationalism thus amounts to a far stronger denial of time than may have been anticipated – it is not just the point structure that is dispensed with, but also our ability to represent any more than one distinct spatial configuration per universe: the interpretation has rendered the formalism dynamically trivial since we can no longer represent change.

A single line of argument is available in defence of total constraint reduction in canonical general relativity against the charge of dynamical trivialisation. Belot (2007, p.78) argues that rather than seeing the reduced phase space as dynamically trivial in the sense outlined above, we should instead reinterpret it as a space of diffeomorphically invariant histories. Thus, we would enable both reductive temporal relationalism and the avoidance of indeterminism but without the cost of trivialising our dynamics. Belot's argument relies on the existence of a canonical isomorphism between the fully reduced phase space and a space of diffeomorphically invariant spacetimes defined via the covariant formalism; we shall therefore dub it the *isomorphism argument*. Because of its importance we shall quote the relevant text in its entirety:

> If one approaches the problem of time via a focus on the transition from the space of initial data to the reduced space of initial data, the problem can appear especially urgent. For in passing from the space of initial data to the reduced space of initial data, one identifies initial data sets that correspond to distinct Cauchy surfaces within a single solution. Prima facie, this involves treating the current state of the universe and its state just after the Big Bang as the same state. Moral: according to general relativity, change is an illusion.
>
> But this is too hasty. For of course the reduced space of initial data is canonically isomorphic to the reduced space of solutions. [55] And in this latter space, some points represent worlds in which there is change (e.g., worlds which begin with a Big Bang) and some represent changeless worlds (e.g., world modelled by Einsteins static solution). So it is hard to see how general relativity teaches us the moral announced.
>
> In the well-behaved theories of section 5 the space of initial data and the space of solutions are symplectically isomorphic, but we nonetheless think of these two spaces as having distinct representational functionsroughly and heuristically speaking, one is suited to represent possible instantaneous states while the other is suited to represent possible worlds. This distinction is grounded by the fact that relative to a slicing one finds that for each $t \in \mathcal{R}$, the map T_{Σ_t} that sends a solution to the initial data that it induces on the instant $\Sigma_t \subset V$ defines a distinct isomorphism between the space of solutions and the space of initial data. This makes it natural to think of points of the latter space as representing states (universals) that can occur at distinct times and to think of points in the space of solutions as representing possible worlds composed out of such states. The elements of this story survived more or less unscathed the introduction of various complicating factors in section 6. But

[55] Belot's Footnote: 'Under the map that sends $[q, \pi]$ to $[g]$ if (q, g) describes the instantaneous state on some Cauchy surface of (V, g).' Where V is a space-time.

> in the case of cosmological general relativity we have only a single canonical isomorphism between the reduced space of initial data and the reduced space of solutions. In this context, it is difficult to deny that the reduced space of solutions and the reduced space of initial data are representationally equivalent. And it seems straightforward that we should interpret points in the reduced space of solutions as representing general relativistic worlds rather than instantaneous states – so we should say that same thing about points in the reduced space of initial data. Thus, we should resist any temptation to think of the reduction procedure as telling us to think of an early state of the universe and a late state of the universe as being the same instantaneous state.

The 'isomorphism argument' contained in this passage essentially runs as follows. Consider a system that does not display any gauge freedom. It's dynamics can be described in terms of a space of solutions to the Euler–Lagrange equations, $\mathcal{S}$, or in terms of set of curves in a phase space, $\mathcal{I}$, with the usual symplectic structure. Although there exists an isomorphism between points in these two spaces, they have distinct *representational roles* – a point in $\gamma \in \mathcal{S}$ represents an entire history of our system, while a point in $\mathcal{I}$ represents an instantaneous state of the system. According to Belot, 'this distinction is grounded by the fact that relative to a slicing one finds that for each $t \in \mathcal{R}$, the map that sends a solution to the initial data that it induces on the instant Σ_t, defines a distinct isomorphism between the space of solutions and the space of initial data.' And thus 'it natural to think of points of the latter space as representing states that can occur at distinct times and to think of points in the space of solutions as representing possible worlds composed out of such states.' If under this interpretation the system is taken to be the whole universe, then clearly points in $\mathcal{I}$ should be considered as representing distinct instantaneous states of the world, and those in $\mathcal{S}$ should be considered as representing worlds composed out of such states.

Now, for a standard gauge theory of the type discussed in Part I, such an interpretation can no longer be justified in these terms. Rather than having a one-to-one map that confers representative equivalence between each time slice of solution and a point in phase space, *for each slice of a given solution* we have a one-to-*many* map, with the target an entire gauge equivalence class of points in phase space. However, if we pass to a reduced phase space $\mathcal{I}_R$ via symplectic reduction as well as constructing a *reduced solution space* $\mathcal{S}_R$ via an analogous reduction process (i.e., quotienting out the action of the Lagrangian gauge group), then we recover our distinct isomorphism per time slice and therefore also our argument towards the different representational roles of the two spaces – the former for instants, the latter for entire worlds.

The case of general relativity – as an atypical gauge theory – is crucially different. Because of the nature of the diffeomorphism group, points in the reduced space of solutions cannot be individually decomposed into slices, and this means that only a single isomorphism exists between each of these diffeomorphically invariant solutions and points in the reduced phase space. This, Belot argues, means 'it is difficult to deny' that we should interpret them as representationally equivalent spaces. Thus, according to Belot *the reduced space is dynamically non-trivial since it can be taken to represent universes that contain evolution* and, furthermore, 'we should resist any temptation to think of the reduction procedure as telling us to think of an early state of the universe and a late state of the universe as being the same instantaneous state' – contrary to what we have argued argued above.

Although innovative and to some extent insightful, the isomorphism argument of Belot is problematic in a number of respects. Firstly, if read as a strong deductive argument, Belot's reasoning seems to rest on the non-sequitur that since the existence of a distinct isomorphism per time slice gives us grounds to fix distinct representational roles for $\mathcal{I}_R$ and $\mathcal{S}_R$, the non-existence of such a family of isomorphisms implies that the two spaces should be taken to be representatively equivalent. Completely besides the nature of the mappings that exist between them, we have very good reasons for asserting that solutions represent worlds and phase space points represent instants – the variational basis upon which the two structures are defined and the different form of the relevant boundary conditions to name just two. Just because in the case of general relativity we no longer have access to one argument towards their representational in-equivalence does not indicate that we no longer have any arguments available at all!

Furthermore, the existence of a single isomorphism between points in two representative spaces is far from a sufficient condition for them to play equivalent roles (although it could in some cases be taken to be necessary) since we can trivially find such relationships between manifestly inequivalent structures – two books with the same number of words, for example. This means that even if we take the isomorphism argument as motivating an interpretation rather than deducing a conclusion, there are good reasons to doubt its strength: without reasons beyond the existence of the isomorphism, *it is not difficult to deny* that the two space are representationally equivalent.

In fact, it makes far more sense for the representational role of a space within a theory to be fixed primarily by its relationship to the representative structures from which it is derived rather than to a space utilised in the context of a different formalism. For the case of general relativity, therefore, it is more appropriate to consider the relationship between the reduced phase space and the unreduced phase space as fixing the former's representational role.

In this context, one could argue that if we accept Belot's interpretation of the reduced space as a space of histories, then we should think ourself forced into also asserting that points in the unreduced space are also representative of four-dimensional histories, and this is manifestly inconsistent with the ADM procedure that leads to the construction of this space. Rather, since we know by definition that a point in the unreduced phase space corresponds to a three metric and its canonical momentum we should take points in the unreduced space to represent instantaneous states and curves in this space to represent entire four-dimensional histories. By passing to a quotient of this space, we are classifying sets of points as equivalent and so *representatively speaking* we are classifying groups of instantaneous states as equivalent. To be consistent with both the representative role of the space from which it is constructed and the manner of its construction, it is difficult to resist the conclusion that the reduced phase space should be interpreted as representing instantaneous states – and therefore that our charge of dynamical triviality against reductive temporal relationalism cannot be avoided.

In any case, as discussed at great length in §10.3, and further analysed at the start of this section, the nature of the Hamiltonian constraint is precisely such that we should *expect* any procedure which treats them as typical, gauge generating constraints to lead to a formalism without nontrivial dynamical evolution. Thus our argument towards fixing the representational roles in the manner described is supplementary rather than fundamental to our conclusion that dynamic trivialisation is implied by reductive temporal relationalism.

In this section we have argued that reductive temporal relationalism and the form of denial of time that it implies is an at best problematic and at worst fatally flawed position. Evidence has been provided that it leads to an interpretation of the formalism of canonical general relativity that is not adequate as a representative framework for describing the world since it admits only static universes. The crucial question is now which aspects of the interpretation are responsible for driving us into such a conceptual *cul-de-sac*? Was it the temporal relationalism or the reductionism that was the cause of the problem? In the next section, we will investigate a different conception of temporal relationalism, in part with the object of settling this matter.

Chapter 12

Denial II: Machian temporal relationalism

12.1 Machian dynamics and the Hamiltonian constraints

A second and quite different perspective on time within general relativity is enabled by the *Machian temporal relationalism* of Barbour (1995, 1994, 2009). The principal element of this form of relationalism with regard to time is not an objection to temporal points forming part of our basic ontology, nor even the assertion of a Leibnizian equivalence principle such that any universes related by temporal symmetries must be judged to constitute the same possibility – although consistency with these other relationalist dictates is implicit. Rather for Barbour the fundamental edict of temporal relationalism is that time should be 'an abstraction, at which we arrive by means of the changes of things; made because we are not restricted to any one definite measure, all being interconnected' (Mach (1960)). This Machian viewpoint on time can be seen as an imperative to try to construct (or at least restructure) our theories in such a way that time does not appear within the basic structure of the theory but is a well defined notion at a derived or emergent concept level. Thus, as well as a position as to what time *is not*, the Machian variant of relationalism is a position as to what time *is*. Particularly, important to both Barbour's interpretation of Mach (which he shares with Mittelstaedt (1976)) and his own philosophy, is that the relational definition of time is a holistic and democratic one based upon contributions from all the motions within the universe. We will take this to mean that a theory or interpretation of a theory that is temporally relational in a Machian sense should provide us with a distinct definition of time for any dynamical history of the universe.

Although it clearly starts from a different perspective, there is a degree of coherence between this form of temporal relationalism and that predicated upon the denial of temporal points discussed above. There is no room within the ontology of a Machian theory for any basic temporal structure since this structure must itself be abstracted out of the ontology and not form part of it. Furthermore, it also seems safe to presume that the democratic nature of the process by which time is abstracted will be such that universes related by temporal symmetries must bear the same emergent notion of temporality. We can therefore expect that even at the level of an abstracted concept of time a Leibnizian equivalence principle should be satisfied.

So much for what it means to be a Machian temporal relationalist in principle. In practice, we have already seen that a formal basis sufficient to establish such a position can easily be achieved within non-relativistic mechanics by switching to the formalism of Jacobi as formulated in §6.2. Again, restating the key ideas for the convenience of the reader. The Jacobi Hamiltonian takes the form

$$H_J = \sum_i p_i.q_i' - L_J = N_J h_J \tag{12.1}$$

where we define the Jacobi Hamiltonian constraint as

$$h_J = \frac{1}{2}\sum_i p_i.p_i + V - E = 0 \tag{12.2}$$

and the lapse is an arbitrary function of of the time label τ. The form of these equations is very suggestive of the canonical formulation of general relativity introduced above. We have a Hamiltonian constraint that is connected with arbitrariness in temporal labelling, and a Hamiltonian that is made up only of a constraint and an arbitrary multiplier. What is particularly important for our purposes is how we should interpret the action of the Jacobi Hamiltonian constraint upon phase space. Explicitly we have that (provided the constraint is satisfied)

$$N_J\{q_i, h_J\} = \frac{\delta q_i}{\delta \tau} \tag{12.3}$$

which indicates that for any specification of the Jacobi lapse the Jacobi Hamiltonian will effect an infinitesimal phase space transformation from the canonical variables characterising a given instant in time to those describing a second instant that is dynamically subsequent.

This is in close analogy to the dynamical role of the Hamiltonian constraints of canonical general relativity. However, as in the relativistic case, this Poisson bracket also encodes a symmetry generating role in that, strictly speaking, the transformation that h_J generates is unphysical because of the dependence on the arbitrary parameterisation encoded in the lapse. Thus again we have evolution enacted by a constraint and thus our dynamics and our temporal symmetry are entangled. In the case of Jacobi's principle, there is a straightforward methodology for disentangling them in the context of Machian temporal relationalism. As mentioned above, as well as the preclusion of external temporal parameters within our mechanical theory, the Machian temporal relationalist position involves a positive idea of time as an equitable measure that can be abstracted from dynamics. Jacobi's principle admits exactly this notion of temporality because we may naturally specify an emergent temporal increment:

$$N\delta\tau = \sqrt{\frac{T}{(E-V)}}d\tau = \sqrt{\frac{\delta q_i.\delta q_i}{2(E-V)}} \tag{12.4}$$

In Lagrangian terms this notion of ephemeris time is introduced by choosing τ such that $T = E - V$. This then leads to the emergent temporal increment $\delta\tau$ and allows us to identify the Hamiltonian lapse with $\sqrt{\frac{T}{(E-V)}}$.[56] The ephemeris time is such that it uniquely and monotonically parameterises dynamical histories. It is holistic and democratic in exactly the sense that Barbour desires because it involves all the dynamical variables of a given system – crucially it is a measure of duration that 'emerges from the dynamics' and 'does not pre-exist in the kinematics' (Barbour (1994, p.2856)).

We can take the Hamiltonian formulation of Jacobi's principle as a model for the Hamiltonian formulation of any Machian temporally relational theory. In particular, it suggests a set of four criteria for the formal structure of such a theory: 1) the parameterisation of phase space curves is arbitrary; 2) the canonical variables do not contain external time variables or their momenta; 3) there is a Hamiltonian constraint that has a dynamical phase space action when combined with an arbitrary multiplier; 4) there exists a methodology for constructing an emergent temporal increment that parameterises dynamical histories in an equitable and unique manner. If 1-4 are satisfied, then it seems reasonable to accept that the theory admits an interpretation consistent with Machian temporal relationalism. As discussed above, such an interpretation has two key features: i) the absence of time in the basic ontological structure and ii) our ability to abstract an equitable measure of duration from the change (or relative change) of the objects that are part of the ontology. Specifically, it seems reasonable to assume that 1-2 lead to i) since they ensure that sequences of points within the phase space can be understood as representing

[56]Thanks to Julian Barbour for clarifying this point to me

the fundamental ontology without reference to time. We then have that 3-4 lead to ii) since, as illustrated by the case of Jacobi's principle, they give us the machinery to associate with pairs of points in the phase space (elements of the ontology) the appropriate temporal increment. The utility of our criteria (which are of a heuristic rather than logical character) is illustrated by their preclusion of a Machian temporal relationalist interpretation of parameterised particle mechanics (there 2 does not hold) and admission of such an interpretation for Barbour–Betotti theory (Barbour and Bertotti (1982)) (where 1-4 all hold).

Given these criteria, we can now address the task of evaluating the interpretation of canonical general relativity in terms of Machian temporal relationalism (MTR).[57] Considering the action (10.3) leads us to conclude that the first and second conditions are satisfied and therefore to expect that aspect i) of MTR holds in canonical general relativity.

To an extent, we also have that the third condition holds because the Hamiltonian is of course a constraint and *in one context* its role is (when combined with the lapse) to generate a transition between dynamically related three geometries. However, as has been asserted throughout our discussion it is essential to remember that that Hamiltonian constraints of canonical gravity have a dual nature with two distinct, context-dependent roles. In the context of a hypersurface embedded within a solution, the role of the Hamiltonian constraints is not of the dynamical type found in Jacobi's theory. Rather, they generate infinitesimal symmetry transformations that form part of the hypersurface deformation group which manifests the fundamental symmetry of the theory. Still, this does not necessarily break the analogy between the relativistic and non-relativistic Hamiltonian constraints since in Jacobi's principle too the Hamiltonian constraint is also connected with unphysical temporal relabellings. However, the fact that the temporal relabellings associated with the infinite set of Hamiltonian constraints of canonical relativity are local and those associated with the single Hamiltonian constraint of Jacobi's principle are global is of crucial importance. Ultimately, the disanalogy that this subtle yet significant difference implies creates an acute problem for an interpretation of canonical general relativity in Machian temporally relational terms.

The fourth criterion that we introduced for the formal structure of MTR Hamiltonian

[57] It must be noted here that much of Barbour's work on the Machian temporal structure of general relativity focuses on general relativity formulated in Lagrangian terms. Our focus on the canonical formalism will not obscure the essential aspects since they are inherent within the dynamical structure of general relativity and therefore beyond the Lagrangian/Hamiltonian distinction. See Pooley (2001) and Butterfield (2002) for detailed philosophical analysis using, for the most part, Barbour's version of the Lagrangian formalism.

theory was that we are able to construct an emergent temporal increment that parameterises dynamical histories in an equitable and unique manner. Given this together with the third criterion, it seems reasonable to presume that we can interpret the phase space of our theory to represent a Machian ontology in the sense of being amenable to the condition ii) above. More explicitly: if our theory is such that two distinct points in phase space that are dynamically related can be connected by the application of the Hamiltonian constraints times suitable multipliers and, furthermore, the difference between them is parameterised uniquely by an emergent time parameter, then we may interpret each point in the phase space as representing the state of the objects in the world and the change between these two distinct ontological states as encoding uniquely a measure of duration in terms of ontological change. In the case of canonical Jacobi's principle, we were able to satisfy this criterion through the employment of ephemeris time, and it is therefore natural to look to construct a similar emergent temporal increment to enable a Machian reading of canonical general relativity.

As has already been mentioned, the crucial difference between the two theories is that the single, global Hamiltonian constraint of Jacobi's principle becomes an infinite set of local Hamiltonian constraints in general relativity. Thus, rather than looking for a single global ephemeris time it is natural to look for many local ephemeris times. A point in the phase space of canonical general relativity corresponds to canonical data on a spacelike hypersurface σ. The locality of the ephemeris times is necessitated by the fact that we need one such time for every $x \in \sigma$. The lapse is of course itself a local function (the dependence on x is suppressed in the notation above). Thus, by looking for formulation of the theory with a lapse such that, in analogy with (12.4), it defines the desired emergent temporal increment when multiplied by an infinitesimal change in the time parameter we can define our local ephemeris times. If we denote the lapse of such a formulation $N_{emph}(x)$, then the local ephemeris can be written simply as $N_{emph}(x)\delta t$, where t is of course now playing the part of the arbitrary time parameter.

The crucial problem is then finding a formulation of the theory containing a local ephemeris with the desired properties. In particular, as well as being insensitive to rescalings of the time parameter, we need our local ephemeris time to be such that it will replicate time as measured by local clocks and thus be consistent with proper time. Furthermore, it is also fundamental to the notion of Machian temporality that any given *local* ephemeris time be an equitable measure of duration, and therefore that it takes account of the contribution of all the other degrees of freedom – even those that are separated from the spatial point at which it is defined.

Interestingly, according to Barbour (2000) the 'deep structure' of general relativity *already* contains exactly the type of local ephemeris time that we are looking for.[58] Starting with the BSW (Baierlein *et al.* (1962)) reformulation of covariant general relativity one can derive Barbour *et al.* (2002, pp.10-12) an expression for the lapse that, within a Lagrangian picture, takes the form $N_{BSW} = \sqrt{T/4R}$ where T is a 'kinetic energy' term Barbour *et al.* (2002, (4.2)). If the time label t within T is chosen such that $N_{BSW} = 1$ then t will correspond to proper time. Furthermore, for arbitrary time label $N_{BSW}(x)\delta t$ *will always be equal to the local proper time*. Just as in the case of Jacobi's theory we can translate this Lagrangian emergent time framework into the a Hamiltonian analogue. There $N_{BSW}(x)\delta t$ gives us an emergent notion of duration that is equal to the local proper time calculated along the direction perpendicular to Λ_t and is non-locally dependent upon the entire three-metric and its canonical momentum. Thus, local ephemeris time *is* a consistent notion within canonical general relativity (given the BSW formulation).

Furthermore, after the introduction of local ephemeris time we are able to classify pairs of points within dynamically successive (infinitesimally close) three-geometries as carrying a trans-temporal notion of identity. Such points are said to be *equilocal*, and the ephemeris time marks them out in terms of the unique temporal metric it provides. For our purposes, the crucial point is that the temporal metric that ephemeris time gives us is defined to be independent of arbitrary reparameterisations of the temporal parameter (t in the case of canonical general relativity) and thus allows us to parameterise dynamical trajectories in phase space in exactly the manner required for criterion 4.

We thus have that 1-4 hold and would then expect canonical general relativity to admit a consistent interpretation in terms of Machian temporal relationalism. However, there is an acute problem with the Machian interpretation resulting, as foreshadowed above, from the locality of the Hamiltonian constraints. The necessary arbitrariness within the definition of $N_{BSW}(x)$ entails that given initial canonical data on a three geometry the dynamical evolution generated by $H(N_{BSW})$ does not provide us with a unique continuation. This is the result of the Hamiltonian's second role of generating infinitesimal diffeomorphisms when considered in the context of hypersurfaces embedded in dynamical spacetimes. Two potential spacetime models are of course represented by a pair of curves within the constraint surface in phase space. Let us assume that these curves are

[58] Note: he does not use the phrase 'deep structure' in this quite this context!

identical up to a given phase space point corresponding to canonical data on σ_0 and thereafter differ only in virtue of a different choice of the lapse – i.e., evolution generated by $H(N'_{BSW})$ rather than $H(N''_{BSW})$. We should then consider them as only differing by a local temporal re-labelling, which (in spacetime terms) can be interpreted as an infinitesimal diffeomorphism of an embedded hypersurface. Given any point $x \in \sigma_0$, there will be an equilocal point within the subsequent three-geometries, σ' and σ'', associated with each of the distinct phase space curves. Thus, we run into exactly the problematic indeterminism discussed in §11.2. This problem does not occur in Jacobi's principle because the arbitrariness that remains within the lapse in that case only gains effect through a single global Hamiltonian constraint and thus cannot lead to distinct phase space curves. Thus, it seems that although an interpretation of canonical general relativity in terms of Machian temporal relationalism can be consistently achieved, this can be done only at the price of admitting ontological indeterminism into a theory that is manifestly deterministic in an empirical and, so far as the conventional interpretation of the covariant formulation goes, formal sense.[59]

12.2 Sophisticated temporal relationalism and indeterminism

Our discussion of §11.1 highlighted the concern that by treating points in the phase space of a gauge theory as representative of individual states we leave ourselves susceptible to a pernicious form of formal indeterminism within a physically deterministic theory. It should be no surprise therefore that, as we have defined it by the criteria 1-4, the Machian temporal relationalist approach to towards the Hamiltonian constraints and phase space of canonical general relativity leads to a specific case of exactly this kind of problem. In our earlier discussion we focused upon symplectic reduction as the supposed remedy for this indeterminism but found that in the case of canonical general relativity such a procedure has a trivialising effect. We are thus in need of an alternative, non-reductive methodology for dealing with indeterminism.

In Chapter 2 we discussed a closely analogous indeterminism issue in the context of phase space. There we considered a strategy for avoiding indeterminism within phase space by identifying gauge related paths as providing representations of the same fundamental history. This identification is made on the basis that the difference between the

[59] In this respect at least our analysis of the Machian viewpoint precisely mirrors that made by Pooley (2001) in the context of the Lagrangian formalism.

histories – when seen as a sequence of ‘objects’ (i.e., instantaneous states) – is merely as to which objects play which roles. If the inflation of possibilities entailed by such haecceitistic differences between distinct histories are discounted – and we therefore adopt a position of histories anti-Haecceitism – then any two histories which are gauge related *in phase space* can be seen as corresponding to a single underlying ontology, and the spectre of pernicious ontological indeterminism vanishes.

An immediate question is then whether adoption of some version of histories anti-Haecceitism can be applied within the context of canonical general relativity, with the object of reliving temporal relationalism of indeterminism. Let us label the combination of histories anti-Haecceitism with relationalism about time as *sophisticated temporal relationalism* (STR). Essentially, a sophisticated temporal relationalist is going to deny the reality of local temporal labellings and endorse the notion that spacetimes related by local relabellings (i.e., re-foliations) are multiply realised in terms of sequences of objects (instantaneous states) that differ merely as to which roles are being played by which objects. In the canonical context, this equates to treating phase space curves that are equivalent to re-foliations of the same spacetime as representing the same fundamental ontology. This is not equivalent to treating the phase space action of the Hamiltonian constraints as generating gauge equivalence classes – such a position is, as we have seen, problematic and manifestly distinct from both the relationalist/substantivalist and (histroies) haecceitist/anti-haeccestist disjuncts. Rather our sophisticated temporal relationalist, unlike the reductive temporal relationalist, can account for the dual role of the Hamiltonian constraints by, on the one hand, treating the curves it generates in phase space as dynamics and, on the other, by classifying the two such curves that are related purely by the deformation of a constituent three-geometry as representing the same basic history realised in terms of two structures that differ merely haecceitistically.

It is very important to note that the identification between the same objects ‘playing different roles’ that grounds the claim that the relevant structures differ ‘merely haecceitistically’ – relies on our ability to consider a pair of three-geometries within space-times related by a re-foliation as being the same object. We thus require a formalism which provides a ‘point-by-point’ identification between each of the constituent ‘objects’ for STR to be a viable position. As was made clear by the analysis of §10.2-3, such an understanding of re-foliation symmetry cannot be provided at the level of phase space. Rather, it is only at the level of hyper-surfaces embedded within a spacetime that we can be the necessary identifications – specifically in terms of the action of an element of the hypersurface deformation group (Teitelboim (1973); Hojman *et al.* (1976)). In that context (or alternatively by using the closely related formalism of Isham and Kuchař (1985b)), we can construct a well defined notion of histories anti-Haecceitism since we can properly define the haecceitistic differences which we which to deny. Thus, our ability to avoid ontological indeterminism *within phase space* is built upon our ability to define the relevant temporal relationalist ontology *within an enlarged space* including embedding variables and therefore inevitably some notion of spacetime .

This last and crucial point makes it clear that the kind of sophisticated temporal relationalist position which we have outlined is going to make an uncomfortable bedfellow for the Machian notion of relationalism. STR as we have defined it essentially makes use of an ontology predicated upon four-dimensional spacetimes and not merely sequences of three-geometries. The relevant inter-structure identification between objects can only be properly defined in the spacetime context. Thus, we violate the key MTR notion that time (or spacetime) should not form part of the basic ontological structure. STR does allow for a viable notion of relationalism (to the extent of allowing us to exclude temporal points) and would seem to be compatible with the emergent notion of time that forms the other key aspect of the Machian position. However, it is essentially a spacetime theory of temporal relationalism and thus cannot be construed as Machian in the most fundamental sense.[60]

12.3 Scale invariance and Machian temporal relationalism

We thus return to the dilemma of extricating the Machian temporal relationalist philosophy from the ontological indeterminism issue. As was mentioned above, the root of the problem lies within the local nature of ephemeris time and this in turn is due to foliation invariance. It is therefore fairly obvious that a solution could lie within the fixing of a foliation and with, therefore, a Machian temporally relationalist interpretation of canonical general relativity in a preferred foliation. Three issues with such a strategy are immediately apparent. First, there must be a basis for this preferred foliation that is, at the very least non-*ad hoc*, and preferably driven by Machian underpinnings. Second, if we are

[60]Here again, we should note a connection between our conclusions, made in a canonical context, and those of Pooley (2001). Although the characterisation given here is different in some notable respects, our STR position is clearly closely related to the 'rather subtle and nebulous form of Machianism' that he defines in terms of a position where one 'regard[s] a spacetime as genuinely constructed from all possible compatible sequences of 3-geometries' (p.17).

to exclude large sectors of the traditional (canonical) solution space by fixing a foliation, then those solutions excluded must be at the very least not empirically grounded, and preferably not empirically viable. Third, the foliation-fixed version of canonical GR must still be consistent with the Machian criteria 1-4 introduced above. Recent years have in fact seen dramatic improvements for the provision of good answers to all three of these points through the development of a scale invariant approach to Machian general relativity know as *shape dynamics*. We do not have space here to give a detailed introduction to or description of this programme and its recent developments. We can at least, however, give a basic outline of its key elements such that we can consider shape dynamics in the context of the three points regarding foliation fixed canonical GR and MTR just raised.

As a philosophical and methodological attitude taken towards physical theory, the Machian approach is one that in general advocates the elimination of absolute or background structure. Modern Machians, Julian Barbour of course being most notable amongst them, argue that whether in Newtonian mechanics or general relativity such structure should be cleaved from our representation of the world via the adoption of alternative, appropriately minimal, theories of mechanics. In this sense, Machianism can be seen as a general scheme for eliminating absolute structure, minimising initial data, and a description of the world based in some sense on relations. This general programme should not be conflated with the specific projects of Machian temporal and spatial relationalism. One would hope, however, that the two cohere – and with regard to absolute structure relating to *scale* and time so it appears to be the case.

There is within all the major theories of mechanics, including general relativity, an absolute notion of scale – conformal transformations (i.e., those which preserve angles but not lengths) are not symmetries at either the local or global level. Within covariant general relativity, this means that solutions of the theory are not invariant under conformal transformations of spacetime. Attempts to construct a gravitational theory that is 4D conformally invariant have a long history stretching back to Weyl (1918, 1922). More pertinent to our project is the programme of constructing a 3D scale invariant theory – i.e., one that is invariant under conformal transformations of space. The investigation of implementing such a symmetry within general relativity in fact parallels the development of the canonical approach in that it can also be traced back to the late fifties and Dirac (1959). In both this work and its extension by York (1973), we already have 3D conformal invariance explicitly connected to a gauge-fixed formulation of general relativity with a preferred foliation. More, recently Gomes *et al.* (2011) have build on the work of Barbour and Ó'Murchadha (2010) and Anderson *et al.* (2005) to propose the existence of a intrinsic *duality* between a theory invariant under volume preserving local 3D conformal transformations and general relativity. Specifically, the particular gauge fixing of general relativity that corresponds to the foliation of spatially compact spacetimes into space-like hypersurfaces of constant mean curvature (the CMC gauge) is taken to be equivalent to a particular gauge fixing of a 'dual theory' that describes sequences of spatial three manifolds invariant under both three-dimensional diffeomorphisms and 3D (volume preserving) conformal transformations. Crudely and yet fairly accurately put, the essence of this shape dynamics programme is then to exchange the foliation symmetry which is present in GR for the local conformal symmetry which is absent. Thus, we can provide a reasoned and, what is more, Machian basis for fixing a foliation.

Our second concern above was that since fixing a foliation amounts to a restriction to a particular sector of the solution space of general relativity there is a danger that it might have undesirable consequences with regard to the empirical adequacy and/or predictive power of the theory. This concern is directly analogous to that discussed in §10.2 regarding the restriction to globally hyperbolic solutions that is entailed by moving to the canonical formalism. In essence, so far as it relates to canonical general relativity, the shape dynamics approach amounts to the introduction of the restriction that as well as being globally hyperbolic solutions must be CMC foliable. According to Gomes *et al.* (2011) this is a weak restriction since it 'includes the vast majority of physically interesting solutions to Einstein's equations while excluding many physically uninteresting solutions'. Thus, one may be able to argue that *empirically* nothing has been lost – certainly we are able to retain the solutions most relevant to currently observed empirical phenomena since the Schwarzschild, FRW, Reissner–Nordström and Kerr–Newman solutions are all CMC foliable (at least so long as we exclude the areas within the event horizon of black hole solutions). Furthermore, as was argued above for the canonical general relativity and the hyperbolic solution case, we are not invoking an ad-hoc philosophical principle in order to exclude these solutions but rather a theory derived from definite physical principles (in this case 3D scale invariance). A more forceful response to this worry is to convert this supposed empirical deficiency into a prediction. Since the restriction to CMC foliable spacetimes can be seen as a consequence of shape dynamics, we may argue that it is providing us with a falsifiable statement about the world that goes beyond those provided by conventional general relativity. Additionally, it also in a sense offers us an *explanation* why our universe does not manifest phenomena relevant to non-

CMC foliable solutions – if they are nomologically possible, why do we not find them or approximations to them in nature? Admittedly, as independent arguments for preferring shape dynamics over traditional general relativity these are not altogether convincing lines of reasoning, but their adoption certainly seems enough to blunt any criticism of the approach along the same lines.

Our third, and most important, worry concerning foliation fixing and MTR is whether general relativity, so formulated, still has the necessary characteristics 1-4 that were deemed necessary for a theory to be susceptible to the relevant Machian relational interpretation of temporality. To investigate this point in the context of shape dynamics, we must consider the latter in a little more technical detail. The methodology for constructing the scale invariant 'dual theory' that Gomes *et al.* (2011) employ can be broken down into five distinct stages. We will briefly outline these in order to argue that the resulting theory can be understood in terms of the notion of Machian temporal relationalism that we have introduced. The first step is to explicitly identify the requisite symmetry that will be *exchanged* for foliation invariance. This is the quotient group denoted by $\mathcal{C}/\mathcal{V}$. Here $\mathcal{C}$ is the (Abelian) group of conformal transformations on the (assumed to be compact) spatial three manifold, which in our notation is σ. The elements of this group are scalars $\phi : \sigma \to \mathbb{R}$ which are such that:

$$q_{ab}(x) \to e^{4\phi x} q_{ab}(x) \tag{12.5}$$

$$P_{ab}(x) \to e^{-4\phi x} P_{ab}(x) \tag{12.6}$$

$\mathcal{V}$ is then a one parameter sub-group representing *homogenous* conformal transformations. The explicit construction of $\mathcal{C}/\mathcal{V}$ in terms of equivalence classes of conformal transformations $[\phi]$ then enforces that there exists a unique representative which leaves the three volume $V_q = \int_\sigma d^3x \sqrt{|q|}(x)$ invariant (see Eq. 61 and the surrounding discussion of Gomes *et al.* (2011) for details). This then allows us to parameterise the group $\mathcal{C}/\mathcal{V}$ by scalars associated with *volume preserving conformal transformations* and thus indicates that we have identified the appropriate symmetry group. The next step is to formally adjoin this symmetry to the theory. Glossing over the technicalities of exactly how this is done (see Gomes *et al.* (2011, §4.1.2)), we can understand this stage in terms of an extension of the phase space of canonical general relativity through the introduction of additional canonical variables (the *Stückelberg field* and its conjugate momenta), which in turn, due to dynamical consistency requirements, results in the presence of an additional set of first class constraints $C(x) = 0$. Like the Hamiltonian constraints, there is one of these constraints per spatial point. However, unlike the Hamiltonian constraints the $C(x)$ *can* be straightforwardly understood as generating unphysical gauge transformations (akin to the transformations generated by the momentum constraints) – these are the volume preserving conformal transformations. Importantly, because of the fact that they are volume preserving one of the new constraints reduces to an identity, so in fact there is one conformal constraint less then there are Hamiltonian constraints. The third step is to impose a gauge fixing via a *best matching* procedure (see Gomes *et al.* (2011, §4.1.3) and references therein) such that all but one of the original Hamiltonian constraints becomes *second class* (in the sense of the standard Dirac (1964) terminology). The usual dynamical consistency conditions of the Dirac prescription for dealing with second class constraints leads to a particular fixing of the lapse up to a one parameter freedom. This lapse fixing is precisely that which gives the equivalence class of CMC foliations. Still following the Dirac procedure, it is possible to eliminate the second class constraints (Gomes *et al.* (2011, §4.1.4)) and arrive at a theory with a Hamiltonian that is constituted by the sum of three distinct types of first class constraint combined with the appropriate multipliers. This new theory is shape dynamics, and its relationship with canonical general relativity is such that for a specific gauge fixing it is equivalent to canonical general relativity in the CMC gauge. Like canonical general relativity, both the symmetry and dynamical properties of the theory are encoded within the structure of the different types of constraints. The first of these constraints are the conformal constraints, which are responsible for the theory's invariance under volume preserving conformal transformations. Next are momentum constraints, which although they have been transformed in the passage to the new phase space can still be understood as implementing three-dimensional diffeomorphism invariance as in the original theory. Finally, and most important for our purposes, there is a *single* Hamiltonian constraint. This constraint is exactly analogous to the single Hamiltonian constraint of Jacobi's principle: it generates dynamics when considered as acting on phase space and global reparameterisations when considered as acting on an entire solution.

Let us now consider our four criteria for a theory to be susceptible to an interpretation in terms of Machian temporal relationalism. Within the dual theory, the parameterisation of phase space curves is arbitrary (i.e., 1), and furthermore the canonical variables do not contain external time variables or their momenta (i.e., 2). We can also now see that, since there is a Hamiltonian constraint that has a dynamical phase space action when combined with an arbitrary multiplier, we also have 3. Thus the condition for an interpretation in terms of MTR is that there exists a methodology for constructing an emergent temporal increment that parameterises dynamical histories in a equitable and unique manner (i.e., 4). Since we have a single Hamiltonian constraint which is combined with a special lapse with a one parameter freedom, intuitively it seems that the construction of the requisite notion of global ephemeris time should be possible within the dual theory itself. Here we will postponing to a future, more detailed, analysis the interesting technical challenge of explicitly constructing such an object (and considering its operational status). Rather, because of the duality between the theories, we can instead simply consider the parallel issue within CMC foliated canonical general relativity. Here it transpires our problem is in fact effectively already solved since it has long been know that all spacetimes admitting a CMC foliation can be parameterised by a unique *geometric time* (See Belot (2007, §7.3) for discussion of the details). Since it is determined by the difference in intrinsic curvature between slices in a dynamical solution this geometric time is both unique and suitably equitable. Thus, almost by definition, canonical general relativity in the CMC gauge satisfies our condition 4. We can therefore assert that both this form of general relativity and the dual theory are amenable to an interpretation in terms of MTR.[61]

[61] Once more there is a close connection between our conclusions and those of Pooley (2001). In essence, his conclusion that the then embryonic 3D conformal theory of Barbour and O'Murchadha (1999) should lead to a solution of Machian indeterminism problem is entirely endorsed by the understanding of shape dynamics which we have presented.

Chapter 13

Denial III: complete observables and the Parmenidean state

We now turn to our third denial of time, which is based on Rovelli's complete and partial observable scheme (Rovelli (1990, 1991, 2002b, 2004)) as applied to canonical general relativity within the work of Dittrich (2006, 2007) and Thiemann (2007). The notion of temporality that is implied by this scheme could be conflated with our first denial in terms of reductive temporal relationalism. However, as we shall see, despite some superficial similarity with regard to how the Hamiltonian constraints are treated there are in fact deep conceptual differences. In particular, whereas reductive relationalism is predicated on the reduced phase space, the Rovelli-Dittrich-Thiemann (RDT) approach is unequivocally non-reductive. Furthermore, whereas reductive temporal relationalism, and for that matter relationalism in general, is fundamentally a thesis with regard to the priority of relational over purely temporal structure, the RDT approach can only be interpreted naturally in terms of a philosophical framework that precludes temporal structure altogether. We will begin our discussion of this third denial by first recalling our earlier treatment (§6.2) of the RDT scheme in the context of the simple nonrelativistic case of Jacobi's principle.

13.1 The complete and partial observables Ansatz

Consider the physical phase space of Jacob's principle, $\Gamma_J = \{(p,q) \in \Pi_J | H(p,q)_J = 0\}$, which is the sub-manifold defined by the satisfaction of the constraint within the full phase space. According to the standard Dirac-Bergmann machinery for dealing with constrained Hamiltonian theories, we define as the *observables* the class of functions on this physical phase space that have vanishing Poisson bracket with the constraints. With the weak inequality implying restriction to the constraint surface, we can write this as a condition on a general phase space function $f : \Gamma_J \to \mathbb{R}$ as $\{f, H_J\} \approx 0$. As with reduction with respect to Hamiltonian constraints in general relativity, the application of this standard definition has immediate, and problematic, consequences for our description of change. If the observable functions must commute with the Hamiltonian, then they must also be non-changing along dynamical trajectories. It seems that either: i) this definition of observable or ii) our expectations for the notion of change that our theory provides, must be adapted to deal with theories in which the Hamiltonian is a constraint. The essence of the RDT approach, both as it applies to Jacobi's principle and general relativity, is to assert that the problem lies within i). With some ingenuity, we can construct observable functions with non-trivial representational capacity so long as we abandon the notion that these obervables change in any conventional sense. The proposal for constructing such obervables is the complete and partial obervables *Ansatz* discussed in §6.2 in terms of the Dittrich (2007) nonrelativistic treatment. It involves us first labelling the configuration variables within phase space *partial observables* and call relations between these variables *complete observables*. The latter are constituted in Jacobi's principle by the reparameterisation invariant specification of the value of one configuration variable with respect to another – as correlations between partial observables. The complete observables are the families of correlation functions that individually give the value of one of the partial observables when the other (the clock variable) is equal to some real number.

Importantly, for a given dynamical solution, the conceptual leeway to consider a *family* of these complete observables rather than a single correlation is dependent on the use of the unreduced formalism. Thus, even though we are in a sense utilising the standard Dirac-Bergmann condition for an observable function we are not thereby committing to the passage to the reduced phase space that is generally assumed to go along with it – we are only being consistent with the Dirac observables scheme so far as it relates to the unreduced phase space. This makes explicit the difference between this approach and reductive relationalism and the first denial of time (Chapter 11). It also implies that, unlike functions of the reduced phase space, complete observables have non-trivial representational capacity since within a given family of observables we may represent the physical structure of a single dynamical universe.

13.2 Application to general relativity

Application of the complete and partial obervables *Ansatz* to canonical general relativity poses a challenge of far greater difficulty for several reasons. We of course have many and not one constraint, and in order to be a true complete observable the object we construct must therefore be constant along the flow associated with all constraints. If all the constraints were mutually Poisson commuting and finite in number, this could be addressed by the technically difficult, but conceptually fairly straightforward, process of: i) introducing one clock variable per constraint and ii) considering as our complete observable a product between each of the flows generated by each of the constraints when applied to a given partial observable, as evaluated for a specific value of each of the relevant flow parameters. We would then have a family of complete observables which are closely analogous to those for Jacobi's principle, only they are now constant along all the various gauge orbits. However, canonical general relativity has of course an infinite number of constraints and, what is more, these constraints do not Poisson commute. As pointed out by Thiemann (2007), even if we restrict ourselves to the space of spatially diffeomorphism-invariant functions (i.e., those satisfying $\{\vec{H}(\vec{N}), f\} = 0$) a flow that is associated with a given Hamiltonian constraint and acts on such a function will not itself be spatially diffeomorphism-invariant since the bracket $\{\vec{H}(\vec{N}), H(N)\} = -\kappa H(\mathfrak{L}_{N_a} N)$ is not invariant. Moreover, even if we remove the momentum constraints altogether and presume ourselves to be working in super-phase-space we still have to deal with the highly non-trivial Poisson bracket between the Hamiltonian constraints, which features structure functions. Thus, the application of the basic RDT scheme outlined above to canonical general relativity poses a significant challenge.

Encouragingly, a number of proposals for meeting this challenge have been put forward. One is that of Dittrich (2006, 2007), which gives an explicit demonstration of how complete observables for general relativity may be constructed in stages by first computing *partially compete observables*. These are complete observables with respect to a sub-algebra of the constraints. One then uses these objects to calculate complete observables with respect to all the constraints. The partial observables in this construction are constituted by spacetime scalars, which in turn are constructed out of canonical fields; this process serves to reduce the number of constraints that must be dealt with. For reasons of space, we will not here attempt an explanation of the details of the Dittrich approach but rather turn our attention to an alternative methodology which makes us of the master constraint programme introduced in §3.2.2. The choice between these two approaches is far from a trivial one and particularly with regard to quantisation it may have significant technical implications. However, for our purposes it must be noted that in respect of the interpretational implications with regard to time the fundamental features are common to both methodologies, and we are choosing to focus on the second only because, given our earlier discussion, it may be introduced more concisely.

Recall that in general the idea of the master constraint programme (Thiemann (2006, 2007)) is to re-write constraint functions, $\varphi_j(p, q) = 0$, in terms of a single equation, which will be satisfied under the same conditions. This new single constraint is then the *master constraint* $\mathbf{M}$. A simple example is given by taking a positive quadratic two form K^{ij} and constructing the equation

$$\mathbf{M} := K^{ij}\varphi_i\varphi_j = 0 \tag{13.1}$$

This equation is satisfied if and only if all the individual constraint functions are vanishing and thus defines the same physical phase space Σ that we had before. A condition for observables on the *extended phase space* is then given by considering the class of functions such that

$$\{\{\mathbf{M}, \mathcal{O}\}, \mathcal{O}\}|_{\mathbf{M}=0} = 0 \tag{13.2}$$

i.e., those functions that have a vanishing double Poisson bracket with the master constraint on the constraint surface. Strictly, this is a restriction that implies that the observable functions generate finite symplectomorphisms that preserve Σ, rather than the usual Dirac-Bergmann condition that the observables are constant along the null directions generated by the individual constraints. However, it can be demonstrated that the two conditions are equivalent (Thiemann (2006)). For canonical general relativity, the explicit form of the master constraint is

$$\mathbf{M} = \frac{1}{2}\int_\sigma d^3x \frac{H(x)^2}{\sqrt{det(q)}(x)} \tag{13.3}$$

This constraint has a number of formal virtues, in particular it is such that its satisfaction implies $H(N) = 0$ for all N, which means that it encodes the same constraint surface as the Hamiltonian constraints. Furthermore, it is also such that $\{\vec{H}(\vec{N}), \mathbf{M}\} = 0$, meaning that it is invariant under spatial diffeomorphisms and will lead us to a constraint algebra with a much simpler form: the *master constraint algebra*, $\mathfrak{M}$:

$$\{H_a(N_a), H_a(N'_a)\} = -\kappa H_a(\mathfrak{L}_{N_a} N'_a) \quad (13.4)$$

$$\{H_a(N_a), \mathbf{M}\} = 0 \quad (13.5)$$

$$\{\mathbf{M}, \mathbf{M}\} = 0 \quad (13.6)$$

We no longer have to deal with the presence of structure functions in our constraint algebra since the highly complex expression (10.10) in the Dirac algebra is replaced by the trivial self-commutation expression (13.6) above. In substituting a single master constraint for the infinite set of Hamiltonian constraints we avoid having to explicitly confront the difficulties of the Poisson bracket algebra with which the latter are associated. Assuming the momentum constraints have been dealt with, either through reduction or via the Dittrich methodology mentioned above, we could now proceed to construct complete observables with respect to single master constraint by considering the flow $\alpha^{\tau}_{\mathbf{M}}$. A family of complete observables is then constituted by the one parameter set of functions defined by the value of one partial observable when the other takes the value s. Assuming these functions are continuous, a given complete observable can then be taken to be invariant under the simultaneous phase space transformations generated by all the Hamiltonian constraints taken together. Thus, as in the case of Jacobi's principle, we arrive at an object which is defined such that it is constant along the dynamical trajectory associated with the relevant 'gauge orbit' – but which has non-trivial representational capacity because it is part of a family of such functions defined within the unreduced formalism. This strange temporal structure is the hallmark of compete observables when applied to the case of Hamiltonian constraints. We now turn to the consideration of the associated interpretational implications for the nature of time.

In our discussion of the Hamiltonian constraints of canonical gravity in §10.3 we emphasised the necessity of treating the constraints such that both the fundamental symmetry of the theory and dynamics are respected. The problem of triviality that beset the reductive temporal relational stance can be understood as a failure on the second count and the problem of indeterminism that troubled the Machian temporal relationalist stance (*sans* a fixed foliation) can be understood as a failure on the first. The kernel of brilliance that allows the RDT scheme to avoid both of these problems is to construct the families of complete observables such that the specification of each family member is deterministic, since they are individually constant along the orbit associated with the Hamiltonian constraints, and yet collectively they are still adequate to represent dynamical universes because of the use of the unreduced phase space. Thus, by endorsing the complete observables as our fundamental object we are provided with an ontology which solves at least one aspect of the problem of time in canonical gravity. However, unlike in a shape dynamics implementation of Machian temporal relationalism, we are not provided with a notion of how to represent change. In the case of our second denial, although time is absent in the sense that it's metrical structure has been relegated to an *emergent* level, it is certainly still a substantive concept since we have temporal ordering of spatial states without our basic ontological structure. Moreover, *change* is still a well defined notion, as we are free to specify the evolution of observable quantities between hypersurfaces with respect to an ephemeris time. What notion of change can we attach to the RDT scheme?

In order to answer this question it is instructive to consider certain key remarks of the three physicists themselves. In discussion of the nonrelativistic application of the complete and partial observables scheme Rovelli (2002b) distinguishes the 'physical phase space' as the 'space of orbits generated by the constraints on the constraint surface' (p.3). In a similar vain Dittrich (2007) defines the physical state as an 'equivalence class of phase space points' which 'can be identified with an n-dimensional gauge orbit' (p1894). So far as they apply to the Hamiltonian constraints as considered acting on the phase space of canonical general relativity such a notion of 'physical space space' and 'physical state' imply an equation between the concept of a history and the concept of a physical state which is radically discontinuous with conventional mechanical theory. Typically states are taken to be instantaneous configurations and histories sequences of such states. In standard gauge theories, where the constraints can be understood unproblematically as generating unphysical transformations, phase space points connected by a gauge orbit are classified as the same state because the difference between them is taken to be unphysical. Dynamical histories are then constituted by either curves within the unreduced phase space which are no-where parallel to these orbits or, more simply by curves, within the reduced phase space. Following the remarks of Rovelli and Dittrich above, the interpretation of change within the complete observables scheme still leads us to classify two points on a 'gauge orbit' as the same state; however this is because the word 'state' is redefined such that in includes all points on the orbit. For the case of the Hamiltonian constraints of general relativity this is simply to adopt a notion of state that involves no temporal specification at all, but rather implies that the observables of a theory are *smeared* everywhere along entire histories. Put more precisely, the complete observables can be understood as 'completely non-local in the unphysical time' (Thiemann (2007, p.81)).

The only viable interpretation of the RDT scheme as applied to general relativity is then one in which time and change have no part – not even at an emergent level. This is to adopt a starkly *Parmenidean* view – time is purely an illusion – and thus constitutes a denial of time in a much stronger sense than that involved in Machian temporal relationalism. It is, however, unlike that involved in reductive temporal relationalism, a denial coherent with the solutions and symmetries fundamental to canonical general relativity. We may still describe dynamically nontrivial universes within the complete observables formalism but almost paradoxically we are able to do this whilst disavowing change. The key to untying this seemingly paradoxical conceptual knot is that, although individual complete observables are eternally frozen, within the families of such observables – which, for a given solution, can only exist because we have avoided reduction – we have access to additional conceptual equipment which allows for the representation of universes corresponding to dynamical spacetimes. Fundamentally, when considered together a family of *doubly* complete observables – constructed by 'smearing out' over the null directions of *both* the Hamiltonian constraints (or Master constraint) and momentum constraints – constitutes a set of *spacetime* correlations in many way analogous to the 'point coincidences' that have variously been proposed to constitute the basic ontology of the covariant formalism.[62] Thus, one must expect that a family of complete observables constructed under the RDT methodology will give us precisely the amount of data needed to reconstruct the 4-D metric tensor for any given (globally hyperbolic) spacetime – including of course those with non-trivial dynamical structure.

Still, one might reasonably raise the question as to in what sense the complete observables are actually observable – clearly they cannot themselves be the *subject* of a measurement as they are entirely non-local! They only feasible way of understanding the relationship between genuine experimental observations/measurements and the complete observables would be to think of a physical measurement to be *constituted* by correlation between various determined values of a variable (i.e., the partial observables). Consider: i) the measurement of a certain variable corresponding to the hand of my watch being in a certain position; and ii) the measurement of a certain variable corresponding to the sun being in certain position – the essence of the complete observables idea is that a genuine

[62]The most famous example of such a coincidence proposal is Einstein (1916) – but also see Westman and Sonego (2008). See (Rickles, 2008, §6.1) for a discussion of the connection between the coincidence type approach and the RDT observables

measurement is just a correlation between i) and ii) defined without reference to any local system of cooridantes. The question still remains whether the observational data customarily associated with general relativity (e.g., deflection of light rays by gravitational bodies) can in practice be reconstructed in these terms. A full analysis would involve considering the construction of compete observables in the presence of matter would therefore go beyond our present 'in vacuo' analysis. The reader is referred to Rovelli (2002a) for a promising line of thought on this front.

More generally, since (as indicated above) families of complete observables could reasonably be understood as the canonical analogues of a set of four-dimensional spacetime correlations defined within the covariant formalism, it seems difficult to consider the problem of reconstructing physical observations in the context of the RDT formalism as any *more* difficult than the problem of constructing physical observations in the context of any covariant scheme based upon spacetime correlations. In essence the compete observables approach is a disavowal of the variation of *all* properties across spacetime and in this sense could be argued to merely be a canonical implementation of one particular interpretation of general covariance.

This brings us to an important qualificatory remark regarding our third denial of time. The Parmenidean position with regard to change that is forced upon us by RDT scheme does not equate to a denial of time either in the sense of asserting that there exists only one time. Nor is it a position that implies that the *temporal dimension* is less fundamental than the spatial dimensions – in of itself it is entirely consistent with a four dimensional spacetime picture of the world. We can therefore see that rather than being allied to the Machian notion of timelessness of our second denial, this third denial is fundamentally antithetical to it. In particular, if we were to couple the application of the complete observables scheme to the Hamiltonian constraints with an application of the scheme to the momentum constraints as well, then the resulting doubly complete observables will be objects smeared non-locally in the unphysical spacetime coordinates and this is an ontology which clearly is not amenable to the Machian temporally relationalist interpretation since it is predicated upon a fundamentally four rather than three dimensional picture of reality. Thus, the choice between our two denials is effectively that between: i) loosing four dimensionality and absolute scale but retaining change; and ii) retaining absolute scale and four dimensionality but loosing change.

Chapter 14

Quantisation and interpretation

At the end of the last chapter we left our discussion at the crossroads between two very different approaches to the interpretation of classical canonical general relativity. What attitude we should take to such scenarios of *metaphysical underdetermination*, both in general and for this specific case, will be the subject of a lengthly analysis in Part IV. Before then we will bring our focus upon the implications of the anti-reductive conclusion of §11.3 to the problem of the quantisation of canonical gravity. The essential question, following on from the discussion of Part II (in particular Chapter 9) and utilising the notion of representational equivalence defined in §4.1, is whether we should understand reduction and quantisation as *commutative* for the case of the Hamiltonian constraints of canonical gravity. If we should, then our various arguments against classical reduction of the constraints can be seen to conceptually undermine any approach which follows the conventional *Dirac route* for the quantisation of gravity. This would certainly be an important result since it would give us cause to question much current research in the field – not least loop quantum gravity as it is currently formulated – and we must therefore proceed carefully. Let us first consider the parallel issue for the momentum constraint.

14.1 Reduction and quantisation of the momentum constraints

Recall from §10.2 that the Poisson brackets between two momentum constraints, (10.8), closes with structure constants. This means that the action of the momentum constraints can be associated with a Lie group. As was mentioned in that earlier discussion, the particular Lie group can be understood explicitly in terms of the implementation of a Lie algebra of diffeomorphisms of the space-like hypersurface σ (Isham and Kuchař (1985a,b)). We would then seem fully justified classically in seeking to: i) quotient the action of these constraints via the application of symplectic reduction; and ii) construct a partially reduced phase space where each point will correspond to canonical variables defined upon a spatially diffeomorphic invariant three geometry. Such a space is the cotangent bundle associated with Wheeler's superspace (Wheeler (1968)) and as such we shall (as above) call it the super-phase-space, $T^{\star}\mathcal{S}$. Formally, its structure is little explored and it is unlikely to be without singularities and other topological complications.[63] However, from a conceptual viewpoint its representational role is clear and we will therefore make the (highly non-trivial) assumption that it has the characteristics of a typical reduced phase space with the associated symplectic geometry. As such, the application of geometric quantisation would be available and a corresponding Hilbert space $\mathcal{H}_{SPS}$ could be constructed.

For our purposes what is most significant is what representational relationship such a Hilbert space would have to that constructed via a Dirac type 'quantise first, reduce second' route. Since the momentum constraints are associated with an algebra which closes with structure constants, it would seem appropriate to think of the associated Lie group as being representable quantum mechanically in terms of the action of a set of unitary operators on an auxiliary Hilbert space. Unfortunately, there is complication here since within modern approaches (i.e., LQG) it is found that we are in fact only able to construct quantum operators generating the finite component of the spatial diffeomorphism group. Although some variant of the group averaging methodology of the RAQ scheme (discussed in §3.2.1 and §4.1) can then be applied,[64] this will lead us ultimately to to a non-separable physical Hilbert space $\mathcal{H}_{mom}$.

The relationship between $\mathcal{H}_{mom}$ and $\mathcal{H}_{SPS}$ (which we would assume to be separable) is not going to be simple. Formally, the two spaces are certainly not going to be unitarily isomorphic and even in terms of the our notion of representational equivalence (as defined by the three criteria introduced in §4.1), we do not have an exact correspondence since the groups involved in the classical and quantum quotienting procedures are strictly speaking different. However, in terms of the putative ontology represented by these two Hilbert spaces, these details are not crucial. The classical and quantum quotients are equivalent in that both lead us to the representation of objects invariant under spatial diffeomorphisms (albeit in slightly different sense since in the former but not in the latter case the diffeomorphisms are smooth). Furthermore, in terms of degrees of freedom we

[63] See Giulini (2009) for detailed discussion of the metric and topological structure of superspace

[64] See Thiemann (2007, §9) for extensive details of such a methodology for the Dirac type quantisation of the momentum constraints in the context of Ashtekar variables.

will have equivalence since in both cases we are cutting down by $6 \times \infty^3$. On the level of observables too we can argue towards equivalence since a representation of an algebra of spatially diffeomorphism invariant observables is well defined on $\mathcal{H}_{mom}$. Clearly, it is reasonable to think of such an algebra as representing the same fundamental objects as the $\hat{\mathcal{O}}_R$ which we would define on $\mathcal{H}_{\mathcal{SPS}}$.

Thus, although we cannot strictly assert representational commutation between reduction and quantisation for the momentum constraints – because of the problems in constructing a quantum operator which generates infinitesimal spatial diffeomorphisms – we can assert commutation to hold *for all intents and purposes* since one may at least represent the same spatially diffeomorphism invariant ontology via both Dirac and reduced quantisation routes.

14.2 How should we interpret the quantum Hamiltonian constraints?

The essential dilemma is whether we should understand the implementation of the Hamiltonian constraints in terms of operators annihilating the wavefunction according to the Dirac quantisation prescription as equivalent to the conceptually problematic classical reduction. Or more precisely, is it appropriate to think of reduction and quantisation as commutative procedures when considered with regard to the Hamiltonian constraints? On a formal level, it is not yet possible to answer this question since the Hamiltonian constraints lie outside the scope of existent commutation proofs. Furthermore, we cannot at the moment even make use of our weaker representative notion of commutativity since we have only established its viability for cases in which the constraints close with structure constants and some variant of the RAQ refinement of Dirac quantisation is available. We can at least argue towards some degree of representative equivalence between the naive quantisation of the Hamiltonian constraint via the original Dirac quantisation methodology (leading to the Wheeler-De Witt equation) and a quantisation of the putative and problematic reduced phase space since there is an equivalence in terms of reduction of degrees of freedom by $2 \times \infty^3$. However, since the original Dirac constraint quantisation methodology does not guarantee us either a well defined physical Hilbert space nor a set of observables and there is no group theoretic basis for interpreting the relevant symmetries, we are still well short of securing even our weak notion of representative equivalence.

Rather, quantisation of the Hamiltonian constraints[65] is the context within which the

[65] We should here, more properly, be speaking of the Hamiltonian constraints as reformulated in Ashtekar

master constraint programme comes into its own. Dittrich and Thiemann (Thiemann (2007, 2006); Dittrich and Thiemann (2006)) have produced encouraging results with regard to the applicability of this scheme to the Hamiltonian constraints (although the significant problem establishing the correct classical limit, among others, still remains) and it would therefore seem, in the first instance, reasonable to assume that if we can establish in general the viability of representational commutation for theories in which MCP has been applied, then we have a good basis for representational commutation in the case of the Hamiltonian constraints.

Recall from above that in the MCP we seek classically to construct a single *master constraint* the satisfaction of which is equivalent to the satisfaction of all the individual constraints. We then promote this single constraint to a self-adjoint operator on an auxiliary Hilbert space and then use the direct integral methodology to construct a well defined physical Hilbert space. To establish representational commutativity we first look to find a correspondence between the classical and quantum reductions in terms of reduction by the same number of degrees of freedom. We can do this be considering the quantum master constraint equation $\hat{\mathbf{M}}\psi = 0$ which we implicitly solve when constructing the physical Hilbert space via the direct integral method. Following Thiemann (2007) we can consider the simple case that $\hat{\mathbf{M}} = K_i \hat{\varphi}_i{}^\dagger \hat{\varphi}_i$ where $K_i > 0$ are constants with the required convergence properties.[66] Next we have that $\hat{\mathbf{M}}\psi = 0$ implies that $\hat{\varphi}_i \psi = 0$ since by definition $\langle \psi \mid \hat{\mathbf{M}}\psi \rangle = K_i \|\hat{\varphi}_i \psi\|^2 = 0$. We can then fall back on the correspondence (Henneaux and Teitelboim (1992)) in terms of reduction in number of degrees of freedom between the Dirac quantum constraint conditions $\hat{\varphi}_i \psi = 0$ and the classical symplectic reduction of a system with physical phase space $\Sigma = \{(p, q) \in \Gamma | \forall_i : \varphi_i(p, q) = 0\}$.

Moving on to the condition regarding observables: we have from §3.2.2 that the MCP allows us to define the strong observables $\hat{\mathcal{O}}_s$ which are such that $[\hat{\mathcal{O}}_s, \hat{\mathbf{M}}] \equiv 0$. What kind of relationship is there between such observables and the $\hat{\mathcal{O}}_R$ that we construct based upon the reduced phase space? We can address this question by first considering the *weak* classical observables which were defined by the double commutator $\{\{\mathbf{M}, \mathcal{O}\}, \mathcal{O}\}|_{\mathbf{M}=0} = 0$. We have that $\{\{\mathbf{M}, \mathcal{O}\}, \mathcal{O}\}|_{\mathbf{M}=0} = 0$ is equivalent to $\{\varphi_i, \mathcal{O}\}|_{\mathbf{M}=0} = 0$. This means that, as noted above, we can think of a geometrical correspondence between $\mathcal{O}$ and $\mathcal{O}_R$ since

variables rather than those expressed in normal ADM variables. However, since the reformulated Hamiltonian constraints close with the same Poisson bracket structure (as they must), this difference is immaterial to our current purpose – although it will become important within a more explicit treatment.

[66] Here the $\hat{\varphi}_i$ are a countable and close-able set of operators which need not be self-adjoint nor form a Lie algebra but are such that $\{0\}$ lies only in their common point spectrum.

the first are constant along the gauge orbits which are quotiented out in order to construct the space in which the latter are defined. Since the classical strong observables (which can be constructed by considering an ergodic mean analogous to (3.5)) are a subset of the weak observables such a correspondence will hold for them also, and it seems correct to think of $\hat{\mathcal{O}}_s$ as being representatively equivalent to a subset of the $\hat{\mathcal{O}}_R$. Classically, we can in fact give a formal criterion to define this subset since they will be such that the pull-back of the map which projects down to the reduced space (i.e., $\pi^\star : \Pi_R \rightarrow \Sigma$) will take them to the $\mathcal{O}_s$. We have not yet considered the representation of the physical state space over which these observables are defined and, as was pointed out above, this relationship is in fact ket to establishing representative correspondence between the observables. We will return to this issue at the end of this section.

More problematic is our condition concerning 'quotienting by the same gauge group' – since MCP is still well defined for cases (such as that of the Hamiltonian constraints) where there is no group theoretic basis to the quotient taken in symplectic reduction, the condition clearly must be adapted to remain relevant. Instead, we should look for the same set of local transformations being removed without any restriction on the nature of these transformations (i.e they may not form a group). Let us return our focus to the Hamiltonian constraints of canonical general relativity. The crucial question is then whether we should understand the MCP as enacting a quantum equivalent of the dynamically trivialising classical reduction discussed in §11.3. In particular, are we doing something equivalent to erroneously treating the (at least) partially dynamical action of the constraints purely as a gauge transformation on the physical phase space? The explicit form of the master constraint for canonical general relativity was given in (13.3). This constraint has a number of formal virtues. In particular it is such that its satisfaction implies that $H(N) = 0$ for all N meaning that encodes the same constraint surface as the Hamiltonian constraints.

Furthermore, it is also such that $\{\vec{H}(\vec{N}), \mathbf{M}\} = 0$ meaning that it is invariant under spatial diffeomorphisms and therefore leads us to a constraint algebra with a much simpler form, the master constraint algebra $\mathfrak{M}$ (13.4-13.6). As was mentioned in our earlier discussion, $\mathfrak{M}$ is by definition such that we no longer have to deal with the presence of structure functions – in substituting a single master constraint for the infinite set of Hamiltonian constraints we avoid having to explicitly confront the complex Bergman-Komar constraint algebra $\mathcal{BK}$ (10.8-10.10). Furthermore, since the master constraint algebra *is* a proper Lie algebra it can be associated with a Lie group of transformations. This means that the task of fully quantising canonical general relativity (i.e., dealing with both sets of constraints) will be made far more tractable.

Returning to the point in hand, clearly $\mathfrak{M} \neq \mathcal{BK}$. So there is a clear sense in which symplectic reduction (which removes the action of the transformations associated with $\mathcal{BK}$) is not going to have a straightforward representational relationship to application of the MCP. Yet, we were able to establish a degree of correspondence in terms of the treatment observables so we should still expect *some* correspondence in terms of which transformations the two reductions treat as unphysical. We might hope to get a definite formal grip on this relationship by calculating the action of $\mathbf{M}$ on a phase space variable. However, since such a calculation will only yield an expression which is vanishing for $\mathbf{M} = 0$ it is clear that the action constructed in this way will be trivial on the physical phase space. The key realisation is that since the Hamilton vector field associated with the master constraint, $X^a_{\mathbf{M}}$, is by definition vanishing on the physical phase space the Poisson bracket between it and any phase space function will always be zero for $\mathbf{M} = 0$. Thus, there are no interpretational difficulties in treating the orbit associated with the integral curves of $X^a_{\mathbf{M}}$ as gauge since it is a trivial move.

To make more definite progress we must consider the quantum theory. Recall from above that we look to represent the master constraint as a positive, self-adjoint operator $\hat{\mathbf{M}}$ on an auxiliary Hilbert space $\mathcal{H}_{aux}$. We then use the direct integral methodology to construct a physical Hilbert space, $\mathcal{H}_{phys}$. Setting aside some important technical complications not least the non-separability of $\mathcal{H}_{aux}$ Thiemann (2007, §10.6.3), the essential elements of this scheme are readily applicable to our master constraint formulation of the Hamiltonian constraints of classical general relativity. What is important for our purpose is whether in constructing $\mathcal{H}_{phys}$ we have carried out a move analogous to treating the classical action of the Hamiltonian constraints on phase space as pure gauge. At first sight, it appears that we have not since the *quantum quotient* that we take in order to construct $\mathcal{H}_{phys}$ is with respect to the kernel of $\hat{\mathbf{M}}$.

Considering things more carefully, the direct integral methodology represents $\hat{\mathbf{M}}$ on $\mathcal{H}^{\oplus}_{aux}(\lambda)$ such that

$$\hat{\mathbf{M}}(\psi_{aux}(\lambda))_{\lambda\in\mathbb{R}} = (\lambda\psi(\lambda))_{\lambda\in\mathbb{R}} \tag{14.1}$$

and then *defines* $\mathcal{H}_{phys}$ in terms of the $\psi_{aux}(\lambda)$ in $\mathcal{H}^{\oplus}_{aux}(\lambda)$ which are such that λ equals zero. This of course means that only states which solve the master constraint will be part of the physical Hilbert space. Furthermore it also means that (following Corichi (2008)) we should think of the quantum equivalent to the Hamilton vector field of the master constraint as vanishing.[67] In fact, since the master constraint can be represented in terms of a positive self adjoint operator on $\mathcal{H}_{aux}$, $\hat{\mathbf{M}}$ is associated with a one parameter family of unitary operators, $\hat{U}(t) = e^{it\hat{\mathbf{M}}}$. It is therefore appropriate to think of the construction of $\mathcal{H}_{phys}$ in terms of the quotienting of a quantum gauge orbit associated with $\hat{U}(t)$ in the same sense as we discussed for the case of RAQ. This would seem to indicate that our intuition from the classical theory has proved correct – quantisation according to the MCP should not, when applied to the Hamiltonian constraints, be consider as involving a *quantum quotienting* analogous to that achieved be reducing out the constraints at a classical level.

We have, however, neglected to consider the observables – it is only in virtue of them that the master constraint can be said to encode the same classical structure as the individual constraints. In fact, according to Thiemann (2006), the requirement that both the observables and the individual constraint operators be represented as self adjoint operators on $\mathcal{H}_{phys}$, can be shown (in solvable models) to fix the inner product such that the solution space must be reduced to the simultaneous one of all constraints. This implies that states in the auxiliary Hilbert space which fail to be solutions of the *individual constraints* will be excluded in the passage to the physical Hilbert space. If this were to hold for the Hamiltonian constraints of canonical general relativity then we would have a restriction on physical states such that they: i) individually solve Wheeler-De Witt type equations of the form $\hat{H}\psi_{phys} = 0$; and ii) collectively solve the master constraint equation $\hat{\mathbf{M}}\psi_{phys} = 0$.[68] Under these circumstances, we can then argue that the representation of physical states arrived at via this 'quantise first, reduce second' methodology will coincide with that based upon quantisation of the dynamically trivial classically reduced space. This is because if the physical Hilbert space is such that only states which are zero eigenvectors of the Hamiltonian constraint operators are permitted, then no two distinct *classical* states which lie along the null direction which the classical constraint function defines can be represented at the quantum level. This means that fundamentally the same set of objects have been excluded from our ontology as in the case of a reduced and then

[67] Our ability to apply these classical geometrical terms in the quantum context derives from the symplectic structure encoded in the space of rays associated with any Hilbert space. See Corichi (2008) and references therein for more details

[68] Whether this proves to be the case in practice can only be established by a full treatment, with the quantum Hamiltonian constraints reformulated in terms of loop variables and the Dirac observables explicitly constructed – presumably using the complete observables Ansatz.

quantised theory. We would therefore be justified in asserting that the quotienting criteria of representational equivalence will hold since we have recovered its fundamental aspect.

Furthermore, this conceptual connection between the physical states also ensures that there is full representative correspondence between the reduced and physical observables at a quantum level and therefore that our criteria concerning observables holds. Thus, for both the general case and the specific case of general relativity the physical Hilbert space constructed via the MCP is representationally equivalent to that based upon quantisation of a reduced phase space – i.e., representational commutation between quantisation and reduction holds. This gives us strong conceptual grounds for doubting the validity of applying this quantisation procedure to the Hamiltonian constraints of general relativity on the grounds of the trivialisation argument of §11.3. Since classically it is incoherent to treat the Hamiltonian constraints as purely generating unphysical phase space transformations, any approach that is equivalent to the implementation of this interpretation at a quantum level will be similarly afflicted.

Chapter 15

Prospectus for a relational quantisation

The arguments of the previous section were aimed at convincing the reader that conventional methodologies for the quantisation of canonical general relativity have dubious conceptual foundations. To a large extent this conclusion is based upon dynamical trivialisation argument concerning reduction and Hamiltonian constraints. Such an argument gives us solid – if indirect – reasons for linking both our viable strategies for understanding classical canonical general relativity to an unconventional approach to quantisation in which the Hamiltonian constraint is not related in the customary Dirac manner leading to a Wheeler-de Witt type equation.

In this respect there is of course much commonality between the main ideas of Part II (the non-relativistic problem of time) and Part III (the relativistic problem of time). Given this similarity, it is tempting to try and reconstruct our arguments of Chapter 9, and in particular consider an analogue of the *relational quantisation*, to the relativistic case. However, for standard canonical general relativity such a proposal is immediately problematic since, unlike in the non-relativistic case, we have an infinite set of *local* Hamiltonian constraints. This blocks a straightforward application of our technique. It does not, however, rule it our altogether. In particular, given either that: i) the infinity of Hamiltonian constraints are rewritten in terms of a single master constraint; or ii) the theory is reformulation of the theory in terms of shape dynamics, we would then regain something like the single global Hamiltonian constraint structure within which relational quantisation has proved applicable. Pursual of the project of formally exploring the first of these ideas will, unfortunately, not be possible within the scope of this current work, and is left to future investigations. However, we can reproduce here some initial work – principally due to Sean Gryb, with the text based upon that of the final section of Gryb and Thébault (2012) – towards the second option.

15.1 Relational quantisation of Shape Dynamics

The Hamiltonian of shape dynamics is given by the sum of three first class constraints H_{gl}, H_a, and C with associated Lagrange multipliers $N(t)$, $N^a(x,t)$, and $\rho(x,t)$ respectively

$$H_{\text{SD}} = N(t)H_{\text{gl}} + \int d^3x \left[N^a(x,t)H_a + \rho(x,t)C\right]. \tag{15.1}$$

Note that the *lapse* $N(t)$ is always homogeneous because the time variable is global. As we saw in §12.3, these constraints can be split into two kinds: i) the constraints that generate gauge transformations and have associated symmetries and ii) the constraint that generates the dynamics. The constraints H_a and C are linear in the momenta and fall under the first kind. We can understand the significance of each by noting the gauge symmetries that they generate. The momentum constraint, H_{a}, is common to both SD and GR and generates infinitesimal spatial diffeomorphisms. The *conformal* constraint, C, generates conformal transformations of the metric of the form

$$g_{ab} \to e^{\phi} g_{ab}. \tag{15.2}$$

These conformal transformations, however, have a global restriction that the total volume of space be preserved. Physically, C requires that the information about the local scale is unphysical. Thus, only angles and ratios of lengths are observable. However, the global scale, set by the spatial volume of the universe, is *not* gauge. This global restriction on scale invariance is crucial because it allows C to be first class with respect to the non-trivial global constraint H_{gl}. In terms of the number of degrees of freedom, this global restriction is also necessary because the two phase space degrees of freedom killed by H_{gl} are recovered by imposing this restriction on C. Thus, the total number of constraints is still equal to that of GR.

The dynamics are generated by the global Hamiltonian constraint H_{gl}. This constraint is *uniquely* defined by the two requirements: i) that the classical dynamics and initial value problem of SD are identical to that of GR; and ii) that it be first class with respect to H_a and C. It is important to point out that the first class requirement implies that H_{gl} is invariant under both spatial diffeomorphisms and conformal transformations that preserve the volume. Unfortunately, H_{gl} is non–local in the sense that it is defined through the formal solution of an elliptic differential equation (given explicitly in Gomes *et al.*

(2011)) which is a modified version of the so–called Lichnerowicz–York equation (York (1972)). It can however, be given explicitly in terms of different perturbative expansions. For our purposes, we will only need the first term of H_{gl} in a large volume expansion. This is a well defined expansion in SD because the volume is a gauge invariant quantity. The details can be found in Gomes *et al.* (2011); Gryb (2011). We will only quote the result:

$$H_{\text{gl}} = 2\Lambda - \frac{3}{8}P^2 + \mathcal{O}(V^{1/3}), \tag{15.3}$$

where Λ is the cosmological constant and P is proportional to the mean of the trace, $\pi^{ab}g_{ab}$, of the metric momenta. For completeness we include its definition (although it will not be used):

$$P = \frac{2}{3}\frac{1}{\int d^3x\sqrt{g}}\int d^3x\pi^{ab}g_{ab}. \tag{15.4}$$

Physically, it is helpful to note that P is the variable canonically conjugate to the spatial volume and is equal to the York time, which is always homogeneous in SD. Note that, to this order in V, the Hamiltonian is homogeneous and leads to the Friedmann universe with pure cosmological constant. Also, in this limit gravity is equivalent to a free particle model like the ones treated earlier in the text, justifying the their use as valid toy models for quantum gravity.

We have now laid out sufficient structure to perform our relational quantisation procedure on SD. We define the central element of the observable algebra, ε, through the Poisson bracket relations

$$\{\varepsilon, g_{ab}\} = \left\{\varepsilon, \pi^{ab}\right\} = 0. \tag{15.5}$$

Its conjugate momentum, τ, is defined by $\{\tau, \varepsilon\} = 0$. We extend the classical phase space to include τ and ε with the Poisson brackets given above and extend the classical Hamiltonian constraint

$$H_{\text{gl}} \to \varepsilon + H_{\text{gl}} \tag{15.6}$$

$$= \varepsilon + 2\Lambda - \frac{3}{8}P^2 + \mathcal{O}(V^{1/3}). \tag{15.7}$$

That this produces an equivalent classical theory can be seen by computing the classical equations of motion for ε

$$\dot{\varepsilon} = \{\varepsilon, NH_{\text{gl}}\} = 0. \tag{15.8}$$

Thus, ε is a constant of motion. We can then integrate out ε in the classical theory and obtain a new Hamiltonian that is just shifted from the original by the constant of motion, $\mathcal{E}$, associated to ε. Clearly, the extension procedure has the effect of redefining the cosmological constant

$$\Lambda \to \Lambda + \frac{1}{2}\mathcal{E}. \tag{15.9}$$

From an operational point of view this requires a change of philosophy: the cosmological constant is seen as a constant of motion rather than a constant of Nature. However, this new interpretation has no effect on the physical predictions of the classical theory.

Despite the fact that the classical theory is unaltered, the quantum theory is noticeably different from that obtained by Dirac quantisation because we require that ε be promoted to an operator. This leads to the following operator constraints on the SD wavefunctional, Ψ,

$$\hat{\varepsilon}\Psi = -i\frac{\partial \Psi}{\partial \tau} = \left(2\Lambda - \frac{3}{8}\hat{P}^2 + \hat{H}_{\mathcal{O}(V^{1/3})}\right)\Psi. \tag{15.10}$$

The cosmological constant can be removed by simply shifting the eigenvalues of the $\hat{\varepsilon}$ operator, just as in the classical theory. We see that the theory we obtain is equivalent to that obtained if we treated the cosmological constant as a global canonical variable rather than a coupling constant. We get a definite time evolution in terms of the global parameter τ.

We can better understand the meaning of this relational quantum theory by considering the nature of the classical intermediary formalism associated with the extend Hamiltonian constraint (15.7). This can be seed to be the SD equivalent to the unimodular gravity theory developed in Brown and York (1989); Henneaux and Teitelboim (1989); Unruh (1989); Unruh and Wald (1989). In particular, in Brown and York (1989), it is shown that promoting the cosmological constant to a canonical variable, in the context of GR, produces a time–dependent quantum theory where the time variable, τ, is canonically conjugate to the cosmological constant. In this case, as in ours, τ is interpreted as the 4–volume of the universe. In GR, the situation is a bit more subtle than in SD because ε is allowed to vary over space. However, as is shown in detail in Henneaux and Teitelboim (1989), there is a secondary constraint $\nabla_a \varepsilon = 0$ that enforces the homogeneity of the ε. Once this constraint is enforced, it is straightforward to see that the Hamiltonian obtained in Henneaux and Teitelboim (1989) is equivalent to the modified SD Hamiltonian (15.7). Thus the relational quantisation of shape dynamics leads to a formalism equivalent to the Dirac quantisation of unimodular shape dynamics. We might, therefore, expect that a prospective relational quantisation of ADM GR would be equivalent to a Dirac quantisation of unimodular gravity. This possibility will be investigated in the following section.

15.2 Relational quantisation of general relativity

The relational quantisation procedure presented in Chapter 9 was motivated by what happens in reparameterisation invariant theories where a single Hamiltonian constraint generates the dynamics. Although the situation is more subtle in GR, where there is a different Hamiltonian constraint for each spatial point, it may still be constructive to check what happens when we apply our quantisation procedure in this case. The GR Hamiltonian can be written as is the sum of two local constraints H and H_a with associated Lagrange multipliers $N(x,t)$, $N^a(x,t)$

$$H_{\text{ADM}} = \int d^3x\sqrt{g}\,(NH + N^a H_a)\,. \tag{15.11}$$

To perform the relational quantisation, we must introduce the central element of the observable algebra ε. However, because the Hamiltonian constraint, H, is a local function of space, so too must be ε. Thus, we must shift H in the following way

$$H(x,t) \to H(x,t) + \varepsilon(x,t), \tag{15.12}$$

where we still have

$$\{\varepsilon, g_{ab}\} = \left\{\varepsilon, \pi^{ab}\right\} = 0. \tag{15.13}$$

The time variable, $\tau(x,t)$, canonically conjugate to $\varepsilon(x,t)$ must also be a local function of space. It would seem that this would produce a qualitatively different theory form the unimodular one previously considered. However, this exact theory has been treated in detail in Gryb (2010). In Section 3.2.4 of that paper, it is shown that the consistency of this theory requires a secondary constraint of of the form $\nabla_a \varepsilon = 0$ and that the resulting theory is identical to the unimodular theory given in Henneaux and Teitelboim (1989). Thus, the relational quantisation procedure (naïvely) applied to GR leads to the standard Dirac quantisation of unimodular gravity.

15.3 Comments on unimodular shape dynamics

Unimodular gravity has been proposed as a possible solution to the problem of time Sorkin (1997). The homogeneous and isotropic case (corresponding to the large volume limit given in Equation (15.3)) has been studied and unitary solutions have been found to exist Daughton *et al.* (1993). Furthermore, it has been argued that treating the cosmological constant as an integration constant rather than a coupling constant could provide a resolution of the cosmological constant problem Smolin (2009). Despite these hopes, there are well–known criticisms for treating unimodular gravity as a genuine solution to the problem of time. These are summarised in Kuchař (1991a). The essential argument is that foliation invariance in GR makes it impossible to genuinely define a global time, which is necessary in the unimodular description. We see these difficulties, in our context, as arising from the fact that our quantisation procedure was designed only to work for theories with a global Hamiltonian. As a result, we can not claim to resolve these difficulties in the context of GR. However, in SD, the situation is considerably improved. In this case, there is a genuine global time parameter and a single Hamiltonian constraint generating dynamics. Thus, the unimodular SD theory presented above is free from the criticisms presented in Kuchař (1991a) and provides a proposal for a genuine solution to the problem of time in quantum gravity.

In essence, our solution is constituted by the application of a three stage procedure: i) translate ADM GR into equivalent shape dynamics formalism; ii) apply extension procedure to construct unimodular shape dynamics; iii) apply standard Dirac quantisation to derive dynamical theory of quantum gravity. Of these three steps, the basis behind the first is perhaps the most contestable; does moving to the shape dynamics formalism not simply amount to sweeping the problem of foliation invariance 'under the rug', rather than solving it? We think not. *On the one hand*, if one considers shape dynamics a fundamental theory of gravity, then we have moved to a formalism that makes manifest a physical deep symmetry triplet of reparameterisation invariance, three dimensional diffeomorphism invariance and three dimensional scale invariance. From this perspective, the issue of retaining foliation invariance within quantum gravity is simply no longer relevant. *On the other hand*, if one insists that general relativity should retain its fundamental status, then – due to duality between that theory and shape dynamics – one can still consider the procedure i-iii above as providing a potential methodology to explore the phenomenology of a foliation invariant theory of quantum gravity not captured within the Wheeler-de Witt formalism. In either case, a quantum theory of unimodular shape dynamics offers an interesting new possibly within the theory space of quantum gravity and warrants consideration of its explicit details, formal consistency and potential for application. Such an investigation will be the subject of future work.

We will return to the consideration of the problem of quantising gravity within the concluding sections of our discussion. In particular, Chapter 20 will examine the conceptual foundations of possible new approaches to the quantisation of gravity in the context of the *underdetermination* precipitated by the two rival formulations of the classical theory corresponding to complete observables and shape dynamics.

Part IV

Realism, Structuralism and Quantisation

Guide to Part IV

Much of the fourth and final part of our discussion will focus on issues from within the philosophy of science and, as such, our theme of the quantisation and interpretation of canonical gravity will for the most part be latent. However, the principal purpose of this lengthly detour will be to provide a philosophical framework within which to analyse situations of *metaphysical underdetermination* – and, as we shall see, such underdetermination can be understood as being embodied precisely in the disparity between the ontology ascribed to canonical general relativity by shape dynamics and the complete observables scheme respectively. Moreover, we will, in the course of our analysis, come to formulate a general philosophical prescription for relating the common structure found in classical underdetermination scenarios to the formal process of quantisation – and it is hoped that this prescription may provide useful insights into the task of quantising gravity.

The philosophy of science discussion of Part IV begins, in Chapter 16, with a number of introductory sections. First, we review the two major frameworks for analysing the structure of a physical theory (§16.1). Next, we consider how one of these frameworks may be used to precisely characterise what it is about a physical theory that could be said to be underdetermined (§16.2). Of particular importance will be the specific case within which the underdetermination is driven by multiplicity within the formalisation of a physical theory. We then introduce the position of scientific realism and explain why one might think it to be specifically threatened by underdetermination cases (§16.3). The next section details the various ways our scientific realist may attempt to break the underdetermination by appeal to external criteria (§16.4), before we introduce the alternative position of ontic structural realism (OSR) within which the ontological bite of the underdetermination is supposedly undercut (§16.5). We will also examine both OSR and scientific realism in the context of the historically grounded undermining of ontology that motivated by the argument from *pessimistic meta-induction*, and from this analysis place a set of conditions on an application of OSR being both consistent and substantive. With these conditions in mind, the final section of this chapter (§16.6) will present a scheme for thinking about formulation underdetermination and OSR in the context of quantisation. The following three chapters will then represent case studies for the analysis of the proceeding ideas within three examples of classical formulation underdetermination. Chapter 17 will examine the Lagrangian and Hamiltonian formulations of Newtonian mechanics, and then Chapter 18 will examine the reduced and unreduced formulations of standard gauge theory, before finally, in Chapter 19, we return our discussion to our two rival formulations of canonical gravity. We conclude, in Chapter 20 with a summary of our project together with an analysis of the relevant implications and prospective research avenues that have been illuminated.

Chapter 16

Metaphysical underdetermination and the interpretation of physical theory

16.1 Theory, interpretation and ontology

Before we embark on the task of investigating the issue of metaphysical underdetermination in the context of the interpretation of physical theory we will briefly consider the two principal frameworks for characterising a physical theory in terms of a formal linguistic system. The choice between these frameworks involves taking a distinct stance as to the structure of physical theories and will, therefore, provide us with a guide as to what we mean when we discuss both the *interpretation* and *ontology* of the theory. Our purpose here will not be to enter into an in-depth discussion of this complex issue. Rather, we merely aim to give an adequate description of the two frameworks and reasonable justification for our adjudication between them based upon the nature of the project in hand.

The *syntactic* framework seeks to provide a representation of the structure of a scientific theories in terms of a formal linguistic system that is interpreted partially by a set of correspondence rules. Following Thompson (1989) (also see Ladyman and Ross (2007) and van Fraassen (1980)) we can unpack these notions into more basic terms. Consider a set of primitive symbols and a set of rules for the formation of formulas using these symbols. A formal language is a set of well formed formulas (wffs) which are defined to be those that contain only primitive symbols (or symbols defined based on primitive symbols) and satisfy the rules of formation. To constitute a formal system we supplement the formal language with two further sets of rules; one which specifies certain wffs as axioms and another that dictates how we can derive the remaining wffs from the axioms. The structures so far defined are sufficient to provide material for the abstract enquiries of logic and pure mathematics but are inadequate for representing empirical science – we need a methodology for providing meaning to the symbols in our language so that it is able to describe the world. In general we designate a provision of meanings to a formal system as an interpretation or model of the system. One way of understanding the syntactic conception of theory structure is in terms of the proposition that a theory can be understood as formal language where the *phenomenal world*[69] is the relevant model.

We can make this idea more precise by first splitting the non-logical symbols of the language into two classes: those that can be assigned meanings directly in terms of observable entities (observational vocabulary); and those that are interpreted in terms of non-observational entities (theoretical vocabulary). If we then assume the existence of a set of rules that provides us with a definition of the theoretical vocabulary in terms of the observational vocabulary then collection of these correspondence rules together with the our two vocabularies defines the phenomenal world as a model for our formal system. In practice useful scientific theories are such that there will always be some terms in the theoretical vocabulary which cannot be reduced to observational terms and thus we have that the correspondence rules only establish a partial interpretation of the language; the observable world is only a partial model. Thus, under the syntactic view we have both that a scientific theory should be constructible in terms of a formal system, a vocabulary of observational and theoretical terms and a set of correspondence rules; and that we should be able to think of the phenomenal world as a model in an appropriate sense.

We now turn to the semantic framework for characterising the structure of a scientific theory. Whereas, in the syntactic conception, a theory is understood as a formal system with the relevant semantics provided by correspondence rules, in the semantic conception these semantics are understood as being provided directly by defining a class of models. In other words, the theory is defined to be the provision of a set of models for a formal system rather than the specification of the phenomenal world as a particular model (or partial model) via correspondence rules. Significantly, such a provision of models, though it implicitly recognises its existence, need not make reference to a formal system at all. Rather, following the *state-space* approach to the semantic characterisation of scientific theory of van Fraassen (1970)[70], we can establish our model class principally by reference to a state space H, each point within which corresponds to a possible configuration of a

[69] We do not here mean to imply that the proponents of the syntactic view are necessarily phenomenalists. 'Phenomenal world' here and below could be understood in any number of a broader sense of observable world

[70] We present van Frassen's view here rather than the set-theoretic alternative because of the natural coherence between it and the state spaces of the Lagrangian/Hamiltonian formalisms.

physical system X that our theory defines. Supplementing this state space by a set of elementary statements E (which are propositions about the magnitude of physically measurable quantities sufficient to characterise the physical system) and a set of satisfaction functions $h(E)$, mapping from the elementary statements to the state space we can form a semi-interpreted language $L =< E, H, h >$. If we define a model M for L in terms of the combination of X and an assignment of a location $loc(X)$ in H to X then all models such that $loc(X) \in h(E)$ will be such that the elementary statements will be true.

Here we have made a distinction between the phenomenal world and an analogous physical system X which represents some aspect of the phenomenal world. Unlike the syntactic view of theory structure the semantic approach treats the latter rather than the former as the proper object of a scientific theory. In this context we can understand the truth of the elementary statements as not being contingent upon actual phenomena rather upon consistency between the system and a model of L. Thus, under the semantic view any well formed physical theory will be true with respect to the system it is concerned with. The connection between this system and the phenomenal world is not taken to be the concern of a different branch of scientific activity distinct from a given theory. We therefore have that under a semantic account two rival theories – one of which may be totally empirically ineffective – will both be true under the lights of the model class and systems that they define. To differentiate between good and bad theories can be taken to principally be to consider a question of empirical adequacy rather than truth. To flesh this idea out we can refine our semantic conception by defining a further element of scientific theory. Van Fraassen presents his picture of the most general features of scientific theories as follows:

> To present a theory is to specify a family of structures, its models; and secondly, to specify certain parts of those models (the empirical substructures) as candidates for the direct representation of observable phenomena. The structures which can be described in experimental and measurement reports we can call appearances: the theory is empirically adequate if it has some model such that all appearances are isomorphic to *empirical substructures* of that model. van Fraassen (1980, p.64) [my italics]

It is crucial, however, to realise that for Van Fraassen the empirical sub-structures do not exhaust the representational capacity of a scientific theory under the semantic conception. Although what *matters* is taken to be empirical adequacy and not the *truth* of how they go beyond the observable phenomena, there is no bar to physical theories describing much more than what is observable. Thus, what will be for our purposes the crucial division within the semantic conception of scientific theories is between the *theoretical structure* of a theory which consists purely of the model class and the *theoretical hypotheses* which detail the representational capacity of the model class with respect to both observable and non-observable entities/objects.

We may now move our discussion on to its principle goal: considering questions of interpretation and ontology in the context of the structure of scientific theory. Let us assume that when we are concerned with questions relating to the *interpretation* of a physical theory we are always principally talking about the manner in which aspects of the formalism can be taken to have a representational capacity with regard to some class of objects or entities. This is reasonable because interpretational disputes with regard to physical theory can usually, if not always, be characterised in these terms – for example, different interpretations of quantum mechanics generally assume the same formalism but take different aspects of it (like the wavefunction) to represent different entities (like information or a guidance wave). The *ontology* of the theory can then be characterised specifically in terms of the provision of an account of what it is that is being represented (i.e., what are the objects or entities).

Under a syntactic conception of scientific theory structure, the division between the observational and theoretical vocabulary defines the interpretation since it is only the observational vocabulary that can straight-forwardly be understood as having representational capacity. The phenomenal world then would seem to define our ontology since it is the thing that, *prima facie*, is being represented. This is a fairly restrictive framework since it does not provide us with an easy means for the characterisation of interpretational disputes concerning the representation of entities that are not directly observable. Furthermore, since our ontology is cashed out in terms of a single model for each theory it is difficult to see how one might deal with theories which only provide us with equivalence classes of solutions. Such disputes and such theories are ubiquitous in the philosophy of physics and although one *might* be able to provide a satisfactory account of them within the syntactic framework, it is certainly not the case that one may do so in an intuitively easy way. An additional worry is that by making such a tight equation between the ontology of a theory and the entire phenomenal world we seem to be divorcing ourselves from the flexibility within real scientific practice. In physics at least, the objects being represented within an individual theory are generally taken to be abstracted away from the phenomenal world itself by use of idealisations or at the very least the demarcation of a certain aspect of the world as 'the system'. Relatedly, since our ontology is, on one understanding, just the phenomenal world itself, *prima facie*, we have little leeway to demarcate ontological attitudes distinct from a simplistic dialectic between realism and scepticism about the phenomena.

The semantic framework on the other hand provides us with adequate conceptual space to carry out a more nuanced exploration of all of these issues. The interpretation of a theory under a semantic conception pertains to the assignment of representational capacity to both observational and non-observational terms (i.e., both to the terms that stand for things like detector readings and the terms that stand for things like the wavefunction). The definition of the theoretical hypotheses of a theory thus allows us room to give empirically equivalent and yet distinct interpretations of a theory. Furthermore, since a theory is now taken to define a class of models we can naturally describe theories which only provide us with and equivalence class of solutions. Finally, with regard to the ontology of a theory the semantic framework provides us with a characterisation which is suitably flexible and amenable to the practice of science. Given an interpretation the ontology can be suitably *unpacked* in terms of the systems that is defined and the observable and non-observable entities which are represented. The relationship between this ontology, on the one hand; and both the phenomenal world and 'reality'[71], on the other, is then manifestly distinguished as a independent issue.

16.2 Underdetermination of what?

The semantic framework for describing the structure of a physical theory provides us with the formalism to easily demarcate three scenarios which can be grouped together under the heading *underdetermination*: theoretical, interpretational and formulation. The first, and most familiar within the philosophy of science, is when we are presented with distinct theories each consistent with the same set of phenomena. In semantic terms this equates either a situation where: i) The same empirical substructures are embeddable within multiple distinct classes of models and therefore we have the same representation of the same phenomena forming part of different theories; or ii) The same phenomena can be represented in terms of *different* empirical substructures which are then in turn embedded in distinct model classes and therefore we have different representations of the same

[71] '[O]ne of the few words which mean nothing without quotes' Nabokov (1955, p.312).

phenomena forming part of different physical theories. We thus, in each case, have an underdetermination of theory by the phenomena such that we have no empirical grounds for deciding between two theories. We also have an underdetermination of ontology since given an interpretation of each of our two competing theories we will then be provided with a choice between two distinct sets of objects or entities (both observable and non-observable) that are taken to exist. It is our attitude towards this ontology that is what, for the most part, will inform our response to this underdetermination. However, before we embark on a discussion of the relation between attitudes to ontology and responses to underdetermination we must first introduce two further notions of underdetermination which are of keener relevance to our discussion of the interpretation of general relativity.

As was discussed above (§15.1), in addition to the division between the empirical and non-empirical aspects of a scientific theory the semantic framework allows for the further division between the model class of a theory (theoretical structure) and the propositions which detail the representational capacity of the model class (the theoretical hypotheses). It is the latter which is specified within the interpretation of a theory and which dictates the class of objects/entities which constitute the ontology. Thus, as well as the underdetermination of ontology that exists between two theories which describe the same phenomena there may be underdetermination of the ontology of a single theory when two or more competing interpretations are available. If the same theoretical structure can be consistently associated with a different set of theoretical hypotheses, then the ontology that is being represented by the theory is underdetermined by our freedom as to its interpretation. The classic example of such an interpretational underdetermination is quantum mechanics where multiple ontologies (e.g., non-local hidden variables or many worlds) may be associated with the same Dirac-Von Neumann mathematical structure via starkly different interpretational stances. Similarly, in quantum field theory it has been noted that one may supplement the same formal structure with an ontology predicated upon a field type or particle type ontology. Once more, the attitude one has towards ontology can be seen to dictate the response one makes to its underdetermination, and in many ways the categories of reaction to interpretational underdetermination parallel those to the underdetermination of theory by phenomena. We have not yet completely exhausted the capacity for scientific theory to underdetermine our ontology. There exists a third, somewhat neglected, and yet pernicious, form of underdetermination that is of most relevance to our project of understanding canonical general relativity.

In addition to the underdetermination entailed by the existence of multiple *interpre-*

tations of a physical theory there is a subtly different class of underdetermination which grows out the existence of multiple *formulations*. In terms of the semantic framework we can understand the formulation of a theory as different ways in which the theoretical structure of a theory (i.e., the model class) can be expressed. The crucial hallmark of distinct formulations as opposed to distinct interpretations is that (as well as being confined to the *structure* aspect of the theory) they are necessarily accompanied with a rigorous translation dictionary which allows us to transform from the language of one formalism to the language of the other. The interpretation and the formulation of a theory are closely related. A given interpretation may make use of a particular formulation of a theory and it may even be the case that a particular formulation is *conducive to* or *exclusive of* a particular interpretation (the nature and relative strength of these notions of formalism underdetermination will be considered more carefully in the next section). Again falling back upon the semantic framework, we can express this interconnection between formalism and interpretation in terms of the limits that differences within the expression of the theoretical structure place upon the construction of the theoretical hypotheses. The strength of the relationship may not be particularly strong, however – such as in the case of quantum mechanics where the various possible formulations (e.g., Schrodinger vs. Heisenberg pictures) are found to licence most, if not all, of the various interpretations equally. However, there is (as the examples that we shall consider illustrate) definite scope for the choice between competing formalisms to be restrictive enough to mandate only certain interpretations and therefore only certain ontologies. Thus, we can encounter underdetermination of ontology which is dictated not by a choice between empirically equivlent theories, nor even between interpretations of the same theory, but rather by the seemingly arbitrary choice between different formulations of the same theory. What attitude we should have to ontology and its underdetermination is the topic to which we now turn.

16.3 The tenets of scientific realism

Realism about the ontology of our best scientific theories is often considered to be the natural or at least the default attitude to take. But what does it entail? Let us quote at length from an influential account Jones (1991, p.185-6):

> Scientific realism is a doctrine about the relationship of our ideas on the nature of things to the nature of things itself. Part of that doctrine is that there is

> a nature of things itself...[Advocates of realism] share the general hope that the scientific enterprise has the capacity to provide accounts of this nature-of-things-itself that are true. In what is more or less the "classical" realist position, this hope is elevated to a belief. Indeed, such classical realists are willing to go out on a limb and claim that theories *in the "mature" areas of science should already be judged as "approximately true"*, and that more recent theories in these areas are closer to the truth than older theories. Classical realists see the more recent theories encompassing the older ones as limiting cases and accounting for such success as they had. These claims are all closely linked to the claim that *the language of entities and processes—both "observational"and "theoretical" ones—in terms of which these theories characterize the-nature-of- things-itself genuinely refers*. That is, there are entities and processes that are part of the nature-of-things-itself that correspond to the ontologies of these theories.
>
> The way in which this reference is fixed, and thus the nature of this correspondence, are topics of intense current debate even among the classical realists who follow the position this far. But their doctrine is a hearty and confident one. It envisions mature science as populating the world with a *clearly defined and described set of objects, properties, and processes*, and progressing by steady refinement of the descriptions and consequent clarification of the referential taxonomy to a full-blown correspondence with the natural order. [my italics]

Realism is here being associated with a number of distinct ideas which come together to form a multifaceted doctrine. We should consider them one by one and see if we can isolate the kernel of the realist position so far as it pertains to the question to the underdetermination of ontology issue. Firstly, according to Jones, the realist (or at least the *classical* realist) believes that particular areas of science are privileged by their maturity and manifest this privilege by the provision of theories which are approximately true. Since we will be dealing exclusively with the unquestionably mature theories of theoretical physics from the nineteenth century and later, this maturity qualification will not be important. Similarly, the relationship between more and less recent theories is not of particular importance since the object of our analysis is not (for the moment at least) underdetermination between historically related theories. What is of significance is the connection drawn between truth (or approximate truth) and the referential relationship between the ontology of a theory and the 'the-nature-of- things-itself'. The nature of this correspondence may be fleshed out in many different ways but what is essential to the realist position is that such a correspondence exists – the ontology vocabulary of a scientific theory genuinely refers. Thus, given the semantic framework for presenting the structure of a theory, the realist can be taken to someone who insist that a strong notion of reference is involved in every viable interpretation of the formal aspect of a theory. The representational capacity of the model class is not merely with respect to the 'ontology' of the theory but with respect to the *Ontology* – by which we mean some notion of 'the-nature-of- things-itself' or 'reality'. It is this second notion of reference that dictates whether a theory carries with it a notion of metaphysical truth as well as the formal truth that is necessarily guaranteed to the models it provides. From a scientific realist perspective it makes sense to say that a given interpretation of a given theoretical formalism can be metaphysically true (or false) in virtue of genuinely referring (or failing to refer) to the Ontology of the world. Significantly, this notion of truth is essentially grounded by a metaphysical rather than empirical or semantic criterion.

In addition to and distinct from this reference tenet Jones involves the realist in the further specification that a scientific theory postulates not a conceptually vague and indefinite ontology but rather one which consists of clearly and distinctly defined sets of objects, properties and processes. Read at face value this would seem a little restrictive in terms of necessarily signing the realist up to non-trivial metaphysical positions with regard to the division of the world into these three categories. We do not need to be so restrictive as to the metaphysical equipment that the realist uses to construct their ontology. Rather, the essence of this second aspect to scientific realism is that the referential relationship that exists between the terms of the ontological vocabulary and their Ontological referents, is one which is between well-defined terms and well-defined entities. Thus, when we find terms such as 'electron' and 'particle' within a theory we should take them to be part of the ontological vocabulary and to refer to electrons and particles.

Intuitively, scientific realism is an attractive stance. In practice it would seem to fit very well with what many scientists, philosophers and children would regard to be the *natural* or *common sense* way of thinking about science and language – there is a sense in which we all behave like realists within our everyday lives. However, the realist view can be seen to be particularly susceptible to being undermined by exactly the type of underdetermination discussed above (§15.2). In general terms, this is because underdetermination scenarios confront us with multiple distinct candidate ontologies and a realist is committed to a referential relationship holding between a single ontology and some real and unique class of entities in the world (the Ontology). Thus, there is an obvious pressure on the realist to either show that true ontological underdetermination cases cannot occur or else do not occur in practice. Alternatively they might wish to show that if underdetermination scenarios do occur we may always break the underdetermination by appeal to external criteria. Above we distinguished three possible levels at which we may have underdetermination: theoretical, interpretational and formalism. We must be careful to distinguish the problem posed to the realist as subtly unique to each type.

The most famous and most discussed notion of underdetermination is between empirically equivalent theories. We can make a useful distinction between *weak* and *strong* empirical underdetermination (this terminology is adapted from Ladyman (2007)). The first (i.e., weak) is when two theories have different possible empirical sub-structures but currently observed phenomena are accounted for indistinguishably as well by either. The second (i.e., strong) results from when either: i) the empirical sub-structures of both theories are isomorphic; *or* ii) both present and all possible future observed phenomena is accounted for indistinguishably as well by the empirical sub-structure of either theory.

Weak underdetermination is of little real concern and is often found in science, in general, and in physics, in particular. The realist, like the practicing scientist, can simply point towards future experiments to probe the phenomenal difference between the two theories and thus effectively collapse this issue into the problem of induction. Given a situation of weak empirical underdetermination between two theories, we can only criticise the insistence that the ontological terms of only one of them genuinely refer to the extent that we endorse inductive scepticism. Every instance of a well confirmed theory in science can be translated into a case of weak empirical underdetermination by adding an inductively sceptical clause that arbitrarily makes a different prediction for future but not past measurements. This is well illustrated by Goodman's (Goodman (1955)) example where we let one theory be any empirical law, such as that all metals expand when heated, and the other be a claim implying that everything observed so far is consistent with the first theory but that the next observation will be different. Finding a response to this kind of weak underdetermination is therefore equivalent to finding a response to the problem of induction. Although this is undoubtably a very important problem it will not serve our discussion to consider any of the vast literature on the matter here and it is enough for our purposes to accept that cases of weak empirical underdetermination are not genuine underdetermination problems *in their own right*.

It is at least arguable that the realist can also side-step strong empirical underdetermination of type ii) on the grounds that it relies on an un-reasonable definition. The idea that theories can have distinct empirical substructures but admit *no possible* phenomenal difference seems to be a dubious assumption. For the first part, it assumes that it is possible to construct a unique notion of what the observable phenomena related to the theory are but such a notion is likely to be interpretation dependent so this seems unlikely. Furthermore, at a given point in time, claiming that there are no possible observable phenomena that will ever be able to distinguish between the distinct structures is effectively to be something like a scientific realist as to theory of observable phenomena. Thus, in order to differentiate weak from strong type ii) empirical underdetermination one must invoke some variant of precisely the doctrine one is attempting to criticise. We therefore have that this notion of empirical underdetermination is also fairly unproblematic for the scientific realist since they seem to have good grounds to simply deny its viability.

In the case of type i) strong empirical underdetermination, however, the challenge cannot be dissolved so easily. That identical or at least isomorphic empirical sub-structures cannot be distinguished between upon the grounds of phenomenal differences seems trivially correct. Furthermore, that such substructures can be embedded within distinct theoretical structures seems consistent with the history of physics at least: the classic example being special relativity and Lorentz's ether theory. Confronted with such cases the realist must either find non-empirical grounds for breaking the underdetermination or else push it down to the next level: what we have here is actually an interpretational underdetermination because any two theories which are such that they are strongly empirically equivalent are actually two interpretations of the same theory. Deferring discussion of the first option to the next section, we now turn to the problem of interpretational underdetermination of ontology which is inherent in recourse to the second.

The difference between two theories and two interpretations of the same theory is, under our semantic conception of a scientific theory, essentially that between two distinct model classes and two distinct sets of propositions which detail the representational capacity of the same model class. Whether the distinction between special relativity and Lorentz's ether theory is more naturally understood in terms the first type of difference or the second is an interesting issue would require a detailed analysis to adjudicate. Perhaps more clear cut is the difference between the various *interpretations* of quantum mechanics. Although arguably some of these may be taken to be distinct theories (dynamical collapse models for example) for the most part the reliance on an identical mathematical formalism should be taken to indicate equivalent semantic structure. It is how that structure is *cashed-out* in terms of an ontology featuring both observable and un-observable entities that is the essence of the distinction. Essentially, interpretational distinctions are grounded in the embedding of the same empirical sub-structure within the same theoretical structure, supplemented by distinct sets of theoretical hypotheses. As such the key issue for the realist is whether these distinct sets of theoretical hypotheses entail distinct casts of ontological entities. If they do, then the realist is faced with a genuine ontological underdetermination case and must recourse to some methodology to privilege one interpretation, and the corresponding ontology, or the other.

It is important to note here that the second aspect of our notion of scientific realism – that which specifies the referential relationship must be between distinct terms and entities – is what blocks the most obvious route of escape from interpretational underdetermination. Since in such cases it is only the theoretical hypotheses and not the theory which is underdetermined, the realist might attempt to claim that they need not adjudicate since they have a single true theory and a single real (and genuinely referential) ontology constituted by the structural overlap between the two sets of interpretationally defined theoretical hypotheses. However, such a stance inevitably involves permitting an ontology constituted by either the shared terms or structural commonalities that preludes at least some of the ontological terms from each interpretation as genuinely referring to distinct existential entities. As such it is not, under our lights, genuine scientific realism but rather some variant of the structural realist programme that will be discussed in §15.5.

Finally, we come to formalism underdetermination. The multiplicitous nature of theoretical formulations has potential be taken to be even more pernicious than that of theoretical interpretation: firstly, this is because of its ubiquitousness in physics theory; secondly, it is because of the insouciance with which it is treated within the practice of that theory. Distinct formulations of Newtonian mechanics date back at least to the late eighteenth century and within relativity theory and quantum mechanics distinct alternatives to the *original* formalism were constructed either simultaneously or fairly shortly afterwards. While the issue of the multiple interpretations of quantum mechanics is regarded as something close to a scandal within physical theory, that one can write Newtons theory of mechanics down in terms of force laws, the Euler-Lagrange equations, Hamiltonian's equations or the Hamiltonian-Jacobi equation is not generally regarded by practicing scientists as matter of concern or even great interest. There is a big difference between multiple formulations of a theory existing and those formulations underdetermining our ontology. The realist would expect that we are *never* dealing with a case of the latter; and in such circumstances the attitude of unconcern with which formal non-uniqueness is treated, is then exactly in line with a pragmatic predilection for realist attitudes towards ontology. Moreover, since two formalisms are, under our definition, necessarily connected by a rigorous translation dictionary, there are grounds for arguing that, unlike in the interpretational case, the ontological differences they may seem to engender will be minimal. The devil of this matter is, as ever, in the detail. In order to establish that the realist ontology can be undetermined by differences that originate at a formal level we will have to carefully consider examples. This is done in Chapters 16, 17 and 18.

On a more abstract level we can consider what criteria must be satisfied for the formalism in particular to underdetermine the realist conception of ontology. The essential point is that the existence of a translation dictionary at the formalism level does not necessitate the existence of a translation dictionary at an interpretation level. Thus, provided it is possible to simultaneously apply distinct interpretations to two formalisms then it is possible that the difference between the two classes of ontological entities that results cannot be reduced to a purely descriptive difference. We can use the example of the Heisenberg and Schrödinger pictures to illustrate this capacity. Labelling these two formalisms $\mathbf{H}$ and $\mathbf{S}$. The key difference between them is that in $\mathbf{H}$ states are represented as static and observables as evolving unitarily, but in $\mathbf{S}$ observables are represented as static and states as evolving unitarily. If the *same* interpretation is applied to both formalisms (for example a simplistic version of the Copenhagen interpretation) then although in a sense we will get distinct ontologies, because of the differing notions of time dependence, this difference can be reduced to a purely descriptive one – the terms *state*, *observable*, *evolution* are used differently but the ontology that is described is the same. To a scientific realist such cases pose no great problem since it is only the description of the ontology that is underdetermined and not the ontology itself – fundamentally we are dealing with the same interpretation, leading to the same class of entities, only it has been applied to two different formalisms. The difference between the two formalisms can be understood in terms of a notational variation and thus does not have any true interpretational weight.[72]

[72]It is important to emphasise here that we are discussing the Schrödinger and Heisenberg *pictures* as different formulations of quantum mechanics and not Schrödinger and Heisenberg *formulations*. The difference between the latter is that between wave and matrix mechanics. As noted by Pooley (2006) this second case may be seen to engender a genuine case of ontological underdetermination since certain interpretations are found to be preferred by certain interpretations.

On the other hand, if different interpretations are applied to different formulations then there is scope for ontology to be underdetermined. However, this is merely a manifestation of interpretational underdetermination discussed above and not a distinct case. The key to genuine cases of formalism/formulation underdetermination (as suggested above) is the possibility of cases where two given formulations of the same theory place different bounds on the cast of viable interpretations. The strength of these bounds demarcates three distinct notions of formalism underdetermination: Firstly, they may be *strict*, meaning that they are such that there is no single interpretation that can be applied to both formalisms. Secondly, they may *exclusive*, meaning that there exists at least one interpretation which is applicable to one formulation but not to another. Thirdly they may be *loose*; meaning that they are such that they make a particular interpretation more natural to one formulation than to another.

All three variants are philosophically interesting and are relevant to the conflict between the tenets of scientific realism and the underdetermination of ontology by physical theory. In the strict case the problem is most acute since there is no unique ontology which the realist has *prima facie* grounds to take to genuinely refer. In the exclusive case the obvious recourse available to the realist is the argument that the interpretations which can be applied to all formalism are privileged over those that are excluded. Reasonable though this may be, it is essentially to bring in an external criteria to break a genuine case of underdetermination and not the dissolution of such a case. As such we will defer discussion of this argument to the next section. Finally, we have the least strong and arguably most prevalent form of ontological underdetermination as mandated by the diversity of formulations; the loose notion. Here the realist might seem to have some grounds to dissolve the seeming formalism underdetermination down to a case of interpretational underdetermination; since it is only whether a particular interpretation is *natural* or not that is in issue, most of the philosophical bite seems to come from the diversity of interpretations not formalisms. However, in practice at least it will prove useful to maintain the distinction between loose formalism and interpretation underdetermination since different strategies of response appear more viable in the one case than the other. In particular, those strategies used to confront strong and exclusive formalism underdetermination are better suited to the loose case than those designed to distinguish between interpretations.

16.4 Breaking the bonds of ontological underdetermination

In each of the three underdetermination scenarios described in the previous section the scientific realist is confronted with at least two distinct ontologies and must then find means to privilege one as true without any empirical grounds for differentiation. There are two distinct ways in which this may be attempted. Firstly, one may argue that, at the level of the ontology itself, for privilege on the grounds of some set of *metaphysical virtues*. Secondly, one may argue at the level of theory, formalism or interpretation for privilege on the grounds of some set of scientific virtues. It is important to note here that the available set of scientific virtues applicable to the particular cases of underdetermination we have in mind is more restrictive than the larger group often invoked to deal with the very common situation of weak empirical underdetermination. Differentiation on the grounds of falsifiability, predictive power or postulation of novel entities is not applicable to cases in which identical empirical sub-structure are involved. Thus, we must look for non-empirical virtues, both scientific and metaphysical.

Primary amongst, and common between, the two sets of non-empirical virtues is the notion of simplicity. In the metaphysical case one might try to argue on the grounds of a parsimony principle that whichever ontology involves a smaller commitment should be that which is endorsed as the true one. However, leaving aside the significant doubt as to whether such a principle can avoid arbitrariness in its specification, it would seem in tension with the tenets of scientific realism that the ultimate grounds for adducting truth should be detached from science in such a way. Thus the attitude of a philosopher like Swinburne (1997) who argues for an 'ultimate a priori epistemic principle that simplicity is evidence for truth' (p.1) does not seem in accordance with the primacy of science as a guide to truth. It is consistent to contend that such an *a priori* simplicity principle is only meant to be applied with all things being equal – thus it may be used as a supplement to, rather than a substitute for, sciences truth determining power. However, the essential point is that the use of any ontological parsimony principle that is justified as *a priori* or by appeal to some wider purely metaphysical principle, is inconsistent with the scientific realists presumed attitude of *scientism* and as such amounts to a revision of their position into an importantly different doctrine.

The same would seem to apply to any other 'free-floating' metaphysical virtue – whether it be consistency with a particular conception of properties and objects or with the fabled analytical philosophers arbiter of *intuition*. We must expect a *scientific* realist to give weight only to metaphysical virtues drawn from a metaphysical framework that is taken to be continuous with physics. However, as pointed out by French (2011), in appealing to virtues drawn from a scientifically loaded metaphysics the realist is threatened by a vicious circularity: it is exactly the underdetermined aspect of the science that is most relevant to the metaphysics. Thus, the scientific realist is forced onto a rack between their scientism and their realism: the realism necessitates a metaphysics that can distinguish between ontologies and the scientism necessitates that this metaphysics must be loaded up with underdetermined science!

We can therefore see why it is natural for a scientific realist to look to scientific and not metaphysical virtues in order to break cases of underdetermination. As mentioned above, the foremost of these is (arguably) again the notion of simplicity. Between theories and between interpretations there is again the danger of arbitrariness. What does it mean for one theory or interpretation of a theory to be simpler than another? We have already mentioned the notion of ontological parsimony but in addition to this we could consider an intra-theoretic or intra-interpretational linguistic notion of simplicity. The key point is that the scientific realist cannot argue for the truth of the simpler scientific theory/interpretation because it is simpler, *simpliciter*. But rather they must base their argument on the notion that simpler theories/interpretation are better because science judges them to be so. Thus, science remains the only arbiter as to truth but it is taken to include simplicity as a virtue amongst its precepts. Whether this is true in practice is a matter of examining the history of scientific practice with an applicable notion of simplicity to hand. This is a huge task not directly relevant to the particular purpose of our discussion of underdetermination.

The major object of our discussion is to frame the particular issue of formalism underdetermination within the context of the wider question of ontological underdetermination With this in mind we can introduce a version of the *simplicity-as-an-underdetermination-breaking-scientific-virtue* argument that is designed specifically for competing formalisms. Following North (2009) we may distinguish a particular aspect of a particular formulation of a physical theory (the example she has in mind and which we shall discuss in detail in the following section is the analytical formulation of Newtonian mechanics) as being in some sense *intrinsic* to the formalism and therefore not dependent on arbitrary aspects of the descriptive apparatus used. An insightful example is the geometric structure of the mathematical spaces used in theories of mechanics:

> The geometric structure of a mathematical space is given by the geometric objects defined on it. Since geometric objects are invariant under coordinate changes, so too is geometric structure. Geometric structure is given by quantities that remain intact while we alter what are merely arbitrary choices of description. This is what we have in the backs of our minds when we say that we are free to choose different coordinate systems for the plane. We mean that choosing different coordinate systems does not alter the underlying structure. It only alters our description of that structure. North (2009, p.6)

The first step in North's argument is the insistence that it is this mathematical structure[73] which should be taken as the ontological vocabulary of a theory of fundamental physics. The application of this *style of interpretation* to different formulations of the same theory is very likely to lead to different notions of ontology. However, once such an ontological underdetermination has been set up exclusively in terms of structure, one then has a precise criteria for differentiating between the two candidate ontologies in terms of simplicity: we simply reject the ontology which is based upon the formulation that uses more than the minimal structure. This argument is immediately applicable to cases in which we are confronted with formulation driven ontological underdetermination of any of our three types since in each case it gives a means to distinguish between any two distinct formalisms. There is also an immediate relevancy and precision to the notion of simplicity utilised. However, again our scientific realist must find scientific grounds for asserting this principle of structural parsimony is a guide to truth. Thus, we must look to the history of science to verify the viability of a scientific realists utilisation of North's arguments (structural minimalism could be used as an *a priori* metaphysical virtue but such an idea would take us away from the focus of our discussion).

The idea that the formulation of a theory with the least structure is that which is most likely to be conducive to future advancement of the field has considerable intuitive currency. However, it in fact has scant support within the modern history of physical theory at least (this is perhaps to be contrasted with the reasonable evidence for simplicity as grounds for success between theories). The major preoccupation of physics during the second half of the twentieth century was with the construction of quantum field theories. Such theories are generally gauge theories and as such contain what Redhead (2003) fa-

[73] n.b. unlike the notion of structure that we will discuss later North's structure *is* particular to a formalism.

mously designated 'surplus structure'. According to North's structural simplicity principle (or at least our interpretation of it) one would expect the development of these theories to be centred upon the non-surplus structure and that all else would have been discarded upon the unceasing march of progression that is modern physics. However, as noted by Redhead and adeptly summarised by French:

> There are numerous examples of the fruitful role of such surplus structure...Redhead himself considered the significance of gauge symmetries within field theory in this context: understanding gauge transformations as acting non-trivially only on the surplus structure, he suggested that non-gauge-invariant properties can enter the theory via this structure leading to further developments via the introduction of yet more surplus structure such as ghost fields, etc. French (2011, p.9)

The role and status of surplus structure within the constrained Hamiltonian formulation of gauge theory will be discussed in more detail in §17.2 so we will not enter in to a detailed analysis of this counter argument to North style structural realism on these grounds here. However, even this brief passage is enough to suggest that there is a good case to be made for surplus structure to be an asset rather than an impediment to scientific development. Thus the argument that we may utilise minimal formal structure as a scientifically informed criteria to distinguish between formalism driven cases of underdetermination of ontology has difficulty getting off the ground.

The potential utility of certain structural aspects of the formulation of a physical theory suggests a different candidate for a scientific virtue that the scientific realist may appeal to for adjudication: heuristic fruitfulness. As in the case of competing theories it is often found – once we are furnished with the admirable vantage of hindsight – that certain formal aspects of one of the candidates have proved invaluable to the successor theories. In the theoretical underdetermination case an obvious example is the notion of Minkowski spacetime which could be taken to lead one naturally from special to general relativity but had no correlate in the scientific *cul-de-sac* that is Lorentzian ether theory. Again bringing our particular focus upon underdetermination between formulations, there seems arguable grounds for, given the appropriate historical evidence, the scientific realist retrospectively privileging that formulation that proved to be more fruitful. We are here, of course, assuming that it is viable for a formulation rather than a theory to have a *successor*; and the extent to which this assumption is a reasonable one can only be establish by historical evidence. However, (among other examples) we are able to point to the example of the Lagrangian formulation of Newtonian mechanics and its successor in Lagrangian quantum field theory to support the notion of a viable formulation-theory successor relationship. It is thus at the very least possible for us to have grounds for privileging that ontology mandated by interpretations of a formalism which has proved to be more heuristically fruitful as being potentially true over and above that mandated by those interpretations that are exclusive to (or at least aligned to) the rival formulation. Thus heuristic fruitfulness may allows us to break the underdetermination so far as it is formulation rather than interpretation driven.

It is crucial, however, that we can form some definite concept about what it means for a formalism in of itself to be heuristically fruitful. This cannot be purely the retrospective specification that, in point of fact, it *has* proved a fertile ground for the genesis of new theoretical structures. Rather we must isolate some feature common to fruitful formulations and be able to utilise this feature to decide between competing formalisms *before* the underdetermination has already been broken by the course of theoretical development. A good candidate for such a feature is captured by the notion of 'heuristic plasticity' (Saunders (1993)) which (according to French) describes the 'feature of certain mathematico-physical entities which allows for their generalisation into new forms, or extension into new domains' French (2011, p.7). Again, we will defer discussion of particular examples to our case studies. But in general, it seems difficult to dispute that the presence of such structures before they have been utilised in a viable successor theory can at most only indicate tentative rather than definite grounds for privilege of one formalism over another – until the arrival of the successor theory it is merely a 'promissory note' for heuristic plasticity. In fact, just as supposedly surplus structure may prove invaluable to scientific development so might seemingly non-plastic structure and what appears to be a heuristically malleable structure according to the presets of one scientific era may be an unfruitful dead end. Despite these concerns the notion of heuristically fruitful structure is too interesting to be discounted altogether and we will return to it in detail when conducting our case studies.

A final virtue, that we will consider in scientific rather than metaphysical terms, and which is particular to interpretational and loose/exclusive formulation underdetermination, is the notion of flexibility. Given a case of underdetermination, where there exist multiple interpretations for a given theory or formalisms that can be given multiple interpretations, we should opt for the formalism/interpretation which is most flexible. Thus, if a certain formalism excludes certain interpretations but another is not so selective then we should favour the more permissive. Similarly, if an interpretation (or family of interpretations) is found to be widely applicable to different theories one might argue that there are grounds for, given a case of underdetermination, that interpretation to be privileged on the grounds of its flexibility. Although tenable, the notion of flexibility as a virtue would seem in conflict with the *explanatory* role of physical theory. If we are always to privilege the most flexible formalism/interpretation then the type of scientific theory that will emerge will be that which places the least restriction on the types of theoretical hypotheses that can be incorporated. However, if we accept the Kitcher (1981) notion of scientific explanation in terms of explanatory unification then this will be to favour an approach to scientific theory which is least explanatory.

This argument can be illustrated explicitly by considering the notions of *argument pattern* and *stringency* that are key to the Kitcher model. Essentially, an argument pattern is an ordered triple consisting of: i) a series of sentences with the non-logical terms replaced by dummy letters; ii) a set of instructions on how to fill these sentences; and iii) a scheme which allows us to classify the sentences as premises or conclusions and which tells us which rules of inference are used. Stringency is then the degree to which a particular argument pattern places restrictions upon the class of arguments that implement it. An explanation is a set of argument patterns which connect a why-question with a class of phenomena and we have that: if an explanation uses a smaller number of more stringent argument patterns to provide a larger the number of conclusions, then it constitutes a more unificatory explanation. This scheme is not the only model for scientific explanation[74]. However, it at least provides us with a good response to the well know problem of irrelevance (Woodward (2009)) which hampers the earlier Deductive-Nomological model (Hempel and Oppenheim (1948)). Furthermore, it also provides us with a notion of what it means for one explanation to be better than another since the provision of a more unificatory explanation intuitively constitutes a deeper and more powerful understanding. This notion of a more explanatory theory however would seem to be in conflict with the preference for a theory including a formalism/interpretation that is more flexible. In particular, preferring more stringent argument patterns is in *prima facie* conflict with preferring theoretical structures which are more flexible.

Thus, we have reasonable grounds to reject this final candidate for a scientific virtue in general terms. Before we embark on the more applied analysis of the three case studies,

[74]The Nerlich (1979) notion of geometric explanation provides an interesting, contrasting example.

we will first investigate an ontological attitude alternative to scientific realism. It will be found that this alternative furnishes us with distinct notions of ontology which drives us to embrace rather than break cases of underdetermination. Thus our test cases will be used as exhibits in the trial not just of scientific realism but of this second viewpoint as well.

16.5 Realism, science and structure

Structuralism as a ontological stance within the philosophy of science has a long history and can be associated with a number of markedly different ontological stances.[75] However, for the purposes of our analysis it will prove instrumental to consider a specific structural realist position that is suitable for the construction of a dialectic with respect to the scientific realist stance. This is the *ontic structural realism* which is defended by Ladyman and Ross (2007), and that arguably draws closest inspiration from the 'best of both worlds' structural realism of Worrall (1989).

The essential difference between ontic structural realism (OSR) and scientific realism is an adjustment within the classification of which aspects of a theory are taken to constitute the ontological vocabulary; along with a corresponding adjustment as to the category of existent *substance* (for want of a better word) which this vocabulary may be taken to refer. Specifically, it is the view that distinct objects or individuals are not fundamental but rather what it real, and what is referred to in the ontological vocabulary of a scientific theory, is inherently structural in nature. As a view it can be distinguished from a more *Aristotelian* flavour of structural realism by the inclusion of mathematical objects such as sets and groups within the class of real structures rather than an emphasis on *concrete structure*.[76] Similarly it is clearly distinguished from an epistemic version of a structural realist stance by being constituted by metaphysical assertions of structural ontology rather than merely the claim that we should only believe in the structural content of theories as an epistemic constraint (Ladyman (1998)).

OSR is advanced by its authors as a solution to a number of thorny problems within the philosophy of science. We shall briefly consider the principal amongst these so as to give some background to the position beyond the underdetermination issue which is our

[75] There is no bar on someone who takes a structuralist view as to scientific theory, simultaneously backing away from realism as to the structural ontology that is being endorsed. Two obvious alternatives are the structural empiricism of van Fraassen, B. C. (2008) and the minimal structuralism (which is more deflationary with regard to the status of individuals) endorsed by Rickles (2008).

[76] Thanks to B. Long for this point.

main focus. The first problem is that of the *pessimistic meta-induction*. Essentially, it is drawn from the observation that historically successful theories are often, if not always, eventually supplanted by successor theories which refer to a distinct or even incommensurable cast of theoretical objects or entities. We can therefore make the inductive argument that, since this has consistently been the case in the past, then it likely to be the case in the present. Thus, provided we accept this *meta-inductive* step, we should also accept that the entities referred to in the (conventionally constituted) ontological vocabulary of our current best theories are very likely to not be referred to by our future best theories. Thus, given the appropriate historical evidence from within mature scientific practice (see Laudan (1981)) we have grounds for doubting the existence of exactly the ontology that the scientific realist wishes to endorse. In essence, the pessimistic meta-induction pulls on the same cord as the underdetermination issue discussed above (§15.2-4) since it serves to cast doubt upon the ability of the scientific realist use our 'best current theories' to give us a notion of ontology (in terms of set of distinct entities at least) that is both robust and unique. Moreover, it also serves to weaken one of the principle *positive* reasons for believing in scientific realism: the viability of explaining the success of science in terms of (approximate) truth and reference.

According to the *no miracles argument* the opponent of scientific realism suffers from a distinct deficit in resources when trying to explain the empirical success of science taken as a whole. While the scientific realist may appeal to the truth tracking nature of the ontological vocabulary with which they endow physical theory as providing an explanation for the past and continued ability of scientific theory to both explain existing phenomena and predicting new phenomena, the anti-realist may make no such appeal. Thus, unless we accept scientific realism, the success of science is adjudged to be miraculous. Putting aside the standard anti-realist responses to such an argument (in particular the provision of a positive, non-realist doctrine for describing both science and its success), it can be argued that the scepticism which pessimistic meta-induction motivates towards the entities involved in the scientific realist stance, serves to undercut the supposedly success-explaining value that these entities have. Essentially both the no miracles argument and pessimistic meta-induction have the same data drawn from this history of scientific theory change, coupled with the improvement of empirical adequacy; and thus, by accepting the first we cast doubt upon the other. In order to motivate realism based upon the history of science we must be provided with an explanatory account of its success that does not fall foul of the meta-inductive existential undermining of the referents of its ontological vocabulary.

In order to meet this challenge within the bounds of scientific realism as we have defined it the most obvious move is to insist that the abandoned terms in the ontological vocabularies of old successful theories did, in fact, genuinely refer by the lights of our current theories. Such an argument can only be truly tested by examination of a stock of historical examples and such an analysis unfortunately falls outside the scope of our current project. However, in general, it seems that we have good reason to be dubious of such a move. Although it may be possible to make the relevant notion of reference rigorous, perhaps by some causal theory of reference (Hardin and Rosenberg (1982)), in removing the entities referred to away from the theories in which the terms doing the referring are defined, we undercut our ability to be genuine scientific realists about our current theories. As pointed out by Ladyman and Ross (2007) such a re-evaluation of reference would imply that Newton was actually talking about geodesic motion in a curved spacetime when he talked about the natural motion of material objects. Meta-inducting from this account of a past successful theory to our current best theories would render the real referents of the ontological vocabulary within these theories as almost completely discontinuous with those same theories! In spirit, if not in essence, this would seem to be contrary to any claim that our best scientific theories are a good description of the world-in-of-itself.

A second, and potentially more successful, methodology for the scientific realist to enable a robust reconstitution of their notion of reference is to adopt the strategy defended by Psillos (1999). We will consider his arguments towards resisting the pull of pessimistic meta-induction in some detail since they arguably constitute the best realist alternative to OSR in dealing with both this and the underdetermination issue. Essential to Psillos' argument is the idea that truly successful scientific theories are distinguished by the provision of (successful) *novel predictions*. He carefully defines this notion in the following terms:

> A 'novel' prediction is typically taken to be the prediction of a phenomenon whose existence is ascertained only after a theory suggests its existence. On this view a prediction counts as novel only if the predicted phenomenon is temporally novel, that is, only if the predicted phenomenon was hitherto unknown...[However,] the notion of novelty should be broader than what is meant by 'temporal novelty'...we should speak of 'use novelty', where, sim-

> ply put, the prediction of a known fact is use-novel relative to a theory, if no information about this phenomenon was used in the construction of the theory which predicted it. Psillos (1999, p.101)

With this, seemingly reasonable, notion of what it means to be be a *genuinely* successful scientific theory in hand one might then set about knocking down many of the historical data points upon which the pessimistic meta-induction argument is based. However, by Psillos' own admission this criterion of success does not serve to exclude all of the relevant historical examples – there are past scientific theories which *did* provide novel predictions despite the fact that they contain theoretical terms within their ontological vocabulary which (by the lights of current theory) *did not* refer. Thus, the notion of novel predictions is not taken to undercut the pessimistic meta-induction on its own. Rather, we make a intra-theoretic division between; those theoretical terms within the ontological vocabulary which are inessential to a theories success in terms of producing novel predictions (the idle terms); and those which are essential to the production of those predictions (the indispensable terms). It is the latter rather than the former which Psillos claims we should be realists about:

> ...the success of past theories did not depend on what we now believe to be fundamentally flawed theoretical claims...the theoretical laws and mechanisms which generated the successes of past theories have been retained in our current scientific image. Psillos (1999, p.104)
>
> ...it is precisely those theoretical constituents which scientists themselves believed to contribute to the successes of their theories (and hence to be supported by the evidence) that tend to get retained in theory change. Whereas, the constituents that do not 'carry-over' tend to be those that scientists themselves considered too speculative and unsupported to be taken seriously...If, therefore, there is a lesson which scientists should teach realists it is that an all-or-nothing realism is not worth fighting for. Psillos (1999, p.107)

For this more selective version of scientific realism to be convincing it must be supported by historical examples that serve to undercut the pessimistic meta-induction; examples where theory change can be described in terms of the discarding of only the predictively idle ontologically relevant terms. This Psillos attempts for the caloric theory of heat and the optical ether theories of the nineteenth century (see his 1999 §6). Putting to one side the strength of his examples (which again we will not examine; see Ladyman and Ross (2007, §2.2.2)), and accepting for the moment that the historical evidence for pessimistic meta-induction can be so undercut; what bearing does this reconstituted scientific realism have on our underdetermination issue? It is certainly conceivable that the disparity between the ontological vocabularies of two empirically underdetermined theories (or interpretations of one theory) might consist of entirely idle terms. In such a scenario we would no longer have a genuine case of ontological underdetermination because the essential ontologically relevant theoretical vocabulary would be common to the two theories (interpretations) and thus can be consistently taken to refer to a unique cast of existent entities. However, the viability of the Psillos version of scientific realism as a response to both underdetermination and pessimistic meta-induction (in general) relies upon *both*: i) our ability to make a precise and principled distinction between the indispensable and idle terms; and ii) this distinction cohering with the terms common between pairs of underdetermined theories (or interpretations) and precursor and successor theories. Over and above the issue of whether his (or other) specific examples may (or may not) be taken to support ii), a strong argument is available against i) based upon an adaptation of the famous *Duhem-Quine Thesis* (DQT).

The DQT (or at least the version of it which we will consider) can be given in terms of two separate sub-theses (Ariew (2011)): i) since empirical statements are interconnected, they cannot be disconfirmed in isolation; and (ii) we can always hold a particular statement true, in-spite of any recalcitrant evidence, by making adjustments to other, not directly empirical, statements (the auxiliary hypotheses) within the theory.

In terms of the semantic conception of theory structure which we have introduced above this (roughly speaking) equates to an ability to always re-embed an empirical sub-structure, incorporating any new phenomena, within essentially the same theoretical structure. This is supposedly always made possible by making small adjustments to the non-fundamental (and therefore auxiliary) aspects to the theoretical structure. This is a particularly strong thesis and if accepted would entail a form of global underdetermination of theories by phenomena. This hardly seems *prima facie* reasonable (close analysis is taken to imply that any form of the DQT will rest on number of highly non-trivial assumptions; see Psillos (1999, p.159) and reference therein for discussion).

However, we need not invoke the full strength version of the DQT to cast doubt upon the viability of Psillos' idle/indispensable distinction. So long as we accept that the terms taken to be essential to a theory's provision of novel predictions have some non-trivial dependence upon at least some of the background theoretical structure used it seems reasonable to posit that the distinction can never me made in absolute terms. In addition to the obvious ambiguity that arises from the type of the dependency between the empirical structures related to the novel predictions and the essential theoretical constituents to which they are taken to correspond, there is also clear scope to argue that the former (and perhaps the latter) will also inevitably be connected to some of the idle terms via the shared background structure. We can formalise our argument as follows (again relying on the semantic conception of scientific theory):

– **P1.** Certain empirical sub-structures within a theory relate to specifically to the novel predictions of the theory

– **P2.** These empirical aspects of the theory are connected to certain theoretical constituents via a set of auxiliary theoretical structures

– **P3.** These theoretical constituents are designated as the 'essential terms', and the rest of the theoretical structure is 'idle'.

I will assume that **P1-3** are uncontroversial. Now, if we then accept that:

– **P4.** There always exists *at least some* alternative auxiliary theoretical structures which lead to a theory with identical predictions but which connect *at least some* of the empirical sub-structures relating specifically to novel predictions to different parts of the theoretical structure.

then we have

– **C.** For any given theory the idle/essential distinction between theoretical aspects of the theoretical structure cannot be made in absolute terms.

Thus, we need only endorse a weakly holistic conception of scientific theory in order to muddy the waters of Psillos' ontologically fundamental distinction. Given that such a notion of science seems, prima facie, irresistible we have good reason to doubt that Psillos' program can succeed since the idle/essential distinction cannot be drawn with sufficient clarity, even in principle.

So, if all-or-nothing scientific realism is not worth flighting for, and Psillos' restrictive version may be taken to be beset by inherent ambiguity, what is there left for a scientific realist to be realist about? What, if any, aspects of scientific theory can survive the gauntlet of pessimistic meta-induction and navigate the ontological undermining of underdetermination? According to the ontic structural realist the only answer is the structure. As mentioned above the inspiration for OSR can be traced to the (non-ontological) structural realist arguments of Worrall (1989). In his influential paper Worrall argues that the force of the pessimistic meta-induction (as embodied by the historical phenomena of scientific revolutions) is strong enough to thwart the type of scientific realist position we have been discussing. However, he also accepts that the argument towards realism based upon the continued novel predictive success of science in general has considerable psychological force:

> The 'no miracles' argument cannot *establish* scientific realism; the claim is only that, other things being equal, a theory's predictive success supplies a *prima facie* plausibility argument in favour of its somehow or other having latched onto the truth.' Worrall (1989, p.102)

Furthermore, he is not satisfied with the standard anti-realist response whereby theories are understood as making no real claims beyond their directly empirical consequences and the continuity of successful empirical content between theories alone is taken to account for science's success:

> Such a [pragmatic or constructive anti-realist] position restores a pleasing, cumulative (or quasi-cumulative) development to science (that is, to the real part of science); but it does so at the expense of sacrificing the no miracles argument entirely. After all, the theoretical science which the pragmatist alleges to be insubstantial and to play a purely codificatory role has, as a matter of fact, often proved fruitful. That is, interpreted literally and therefore treated as claims about the structure of the world, theories have yielded testable consequences over and above those they were introduced to codify and those consequences have turned out to be correct when checked empirically. Why? The pragmatist asserts that there is no answer. Worrall (1989, p.102)

The key point taken by Ladyman and Ross (2007) from Worrall's dual use of the no-miracles argument *and* pessimistic meta-induction is that there is considerable force pushing us towards a view in which there is something objectively real corresponding to science's description of the world (and therefore no need for miracles) but this something is not the distinct cast of entities and objects that are undermined by pessimistic meta-induction. Rather, according to Ladyman and Ross, we should avoid being 'metaphysically committed to the existence of self-subsistent individuals' and assert that what exists are 'real patterns' which are the referents of the structural components of scientific theory. Furthermore, they also assert that the 'material mode' of reality additionally contains objective modal structures which are in turn represented by logico-mathematical formal modal structures. These structures are taken to be retained and developed between successive scientific theories over and above dramatic change in the theoretical entities described by the theories and thus, by rejecting the conventional realist specification of the ontological vocabulary, the proponent of OSR is able to undercut the anti-realist thrust of the pessimistic meta-induction. Simultaneously, by providing a realist type account of science with a robust referential relationship holding between the relevant structural aspects of theory and reality, OSR enables a stance which can ride the psychological current of the no miracles argument:

> If theorists are able sometimes to capture the objective modal structure of the world then it is no surprise that successful novel prediction sometimes works, and the practice of theory conjunction ought to lead to progress at the empirical level (Ladyman and Ross 2007 p. 123)

Thus, as a position taken specifically with regard to problems deriving from tension between our philosophical description of science and the history of science's development and success, there are good reasons to favour the ontic structural realist programme.

The programme also has potent applications within modern theories of physics where the notion of primitive self-subsisting individuals/objects is adjudged to be problematic (or at the very least be underdetermined by the interpretation of the theory in question). French (2011), notes that the *received* view on quantum statistics (both the Bose-Einstein and Fermi-Dirac types) is that they imply that particles can no longer be regarded as individuals. In a similar vein to this received view, Redhead (1999) argues that because the global number operator of a relativistic quantum field theory cannot be broken down such that it gives us a unique local notion of particle number:

> In a truly local physics, particles don't exist in the relativistic theories, except in an attenuated 'approximate' sense, which may be good enough for physicists but definitely not for philosophers trying to understand in absolute 'all or nothing' categories what QFT is about!

Teller (1995) makes use of similar arguments to motivate an ontology of 'quanta' rather than particles possessing (self-subsisting) primitive identity within QFT. However, with regard to the non-relativistic quantum mechanical case at least, the case against individuals is not entirely straightforward since the novelties of quantum statistics may be explained in terms of restriction on allowed states of a quantum system rather than a indication of the absence of primitive individuals:

> With the reduction in statistical weight explained by the inaccessibility of certain states, rather than by the non-classical metaphysical nature of the particles as non-individuals, one can continue to regard them as individuals for which certain states are now inaccessible – just because the particle labels are statistically otiose does not mean they are metaphysically so. French and Krause (2006, p.150)

Thus, we can conceive of the situation as one of (in our terms) interpretational underdetermination of the ontology of quantum mechanics: two interpretation packages are available for the same quantum mechanical formalism, *particles-as-non-individuals* and *particles-as-individuals* (subject to some accessibility constraints). One might therefore reasonably conceive of the situation in terms of these two competing interpretations of individuality within quantum mechanics as simply paralleling the competing realist interpretations of science as embodied by OSR and the 'object-orientated' standard scientific realism of a philosopher such as Psillos. However, the ontic structural realist can avail themselves of a more sophisticated response to this underdetermination which is unavailable to the traditional scientific realist. Following Ladyman (1998) an advocate of OSR may argue that both packages are merely manifestations of the underlying purely structural ontology. In practice, for this case it means that we arrive at a picture such that:

> ...a particle will be understood as a fermion, say, in terms of the relevant (anti-symmetric) representation of the permutation group (and hence the relevant symmetry of the wave-function) and as an electron in terms of properties of mass and spin associated with the relevant irreducible representation of the Poincaré group, and so on. French and Krause (2006, p.173)

At this point a clear objection that can be made is that the particular variant of ontological structural realism that Ladyman, Ross and French are advocating is not the only option. We could perhaps be deflationary rather than eliminative towards individuals/objects and adopt something closer to Rickles' (2007) minimal structuralism. In fact, it could seem in general that the OSR programme implicates itself in a response to situations of ontological underdetermination within science which is to itself underdetermined. Such criticisms do seem well placed but the subtleties which their evaluation requires are not of direct relevance to our discussion. In particular, to the extent to which a less metaphysically loaded structuralism provides a satisfactory resolution of issues of underdetermination *in virtue of its structuralism*, the OSR programme will also provide a satisfactory resolution. Moreover, the dialectic with the scientific realist, which is fundamental to our analysis, is most clearly constituted using the *more metaphysical* variant of structuralism that we have been considering precisely because of the *thicker* notion of ontology of entities/objects which it entails.

The most pertinent feature of OSR for our current discussion is the extent to which its structural notion of ontology allows for the dissolution of cases of formalism underdetermination without the abandonment of either the realism or the science aspect of scientific realism. Essentially, we can understand the scientific realist as committed to three key ideas: i) the fundamental supremacy of (mature) science as a guide to the nature of reality; ii) the genuine referential relationship existing between the ontological vocabulary of our best scientific theories and world-in-of-itself; iii) the constitution of both the ontological vocabulary and that to which it referees in terms of distinct classes of objects and entities. As has already been mentioned above, OSR involves the rejection of the third of these and, on a superficial level at least, we can see that it is this move that engenders a resolution of the underdetermination issue without recourse to anti-realism. Given two formulations of a theory interpreted in terms of two different sets of entities we can avoid underdetermination by constituting our ontology not by the entities themselves but by some overall structure lying behind them. It is this structure that is real and which we take to be referred to by the structural ontological vocabulary of the theory.

A well placed criticism (due to Pooley (2006)) that has been levelled at OSR, is that the specification of a structural ontology may not on its own be sufficient to resolve a genuine case of ontological underdetermination, at least as it exists between formulations. As is well illustrated by the type of 'structural realism' defended by North, it is quite possible for the structural ontological vocabulary itself to be underdetermined if it is characterised in such a way as to be particular to each formalism (see the three cases studies of Chapters 16, 17 and 18 on this point). Thus, the structures that the defenders of OSR are looking to endorse must be such that they span between the appropriate formulations – it must be common structure. An obvious candidate for such structure is the mathematical transformations and interrelations that constitute the translation dictionary between two formulations. However, such a characterisation of the structural ontology is also problematic. As further noted by Pooley (2006), such interrelations between formulations offer only a very thin notion of structure that alone seems insufficient to be the fundamental furniture of the world: what is needed is a 'single, unifying framework [which we can] interpret as corresponding more faithfully to reality than do its various [conventional] realist representations' (p.7). Thus, the onus is on the ontic structural realist to offer more than a purely set or group theoretic *characterisation* of the common underlying structure invoked to dissolve cases of underdetermination. What is needed, in essence, is a physico-mathematical framework that generalises the structures relevant to each formulation in such a way as to illustrate that each formulation is merely a different representation of the same underlying 'reality'. Such a framework must inevitably include dynamical[77] as well as purely mathematical aspects and will therefore be particular to the formulations and theories to which it pertains. The extent to which this highly nontrivial task proves possible in practice will be one of the key issues examined via the investigation of our three test cases in the following three chapters.

If we, for the time being, accept that ontic structural realism *does* provide a good solution to problems of ontological underdetermination. This would mean that in addition to its utility for providing a notion of ontology that evades both the no-miracles and pessimistic meta-inductions problems, OSR would have the additional strength of providing us a solution to the metaphysical underdetermination issue. Importantly, there is no guarantee that the notion of structural ontology particular to the solution of the two different problem types will cohere. It seems perfectly possible that the structural commonalities that are retained between precursor and successor theories will be found to be of an entirely different type to those held, for example, to solve the quantum individuals interpretational underdetermination issue. In that case, at least, one could argue that exactly the same group theoretic structure invoked by French to resolve the underdetermination of interpretation issue, could be identified as the 'common structure' between quantum mechanics and its classical mechanical forebears. So there is a *prima facie*, viability to such consistency existing. However, as we have just been discussing, mere group theoretic mathematical structure might not seem sufficient to constitute our ontology. In general, the task of providing the kind of generalising physico-mathematical framework that is

[77] On this point see Bain (2009).

taken to be required for dealing with underdetermination issues *within* a theory seems difficult enough without the additional requirement that this framework also be common *between* theories. To some extent the two tasks seem to pull in different directions since at least under a revolutionary type conception of theory change we might expect the notion of a generalising framework between successive theories to be very hard to construct.

This question of self-consistency within a structural view of scientific ontology will be taken up in the next section in the specific context of formalism underdetermination within a classical theory and historical inter-theoretic structural continuity with the corresponding quantum theories. Our major focus will be whether of not structures can be isolated that are simultaneously: i) common between formalisms; ii) not undercut by PMI; iii) is a genuine physico-mathematical framework that includes dynamical structure. We will thus not here make a detailed consideration of whether we could hope to meet Pooley's challenge that a viable structural realist ontology must be able to be understood 'as corresponding faithfully to reality', in the same sense as a traditional realist ontology. Answering such a question relies upon a subtle understanding of what attitude towards ontology the ontic structural realist may take and, as discussed above, such issues will be neglected in our current discussion.

16.6 Quantisation, structuralism and underdetermination

We will here give a brief recapitulation of some of the principal arguments and concepts that have already been introduced within earlier parts of this thesis. We do this both for the sake of clarity, and in order to further motivate the analysis of the next section. Let us start with quantisation. Within Chapter 3 we detailed three methodologies for the quantisation of a classical standard gauge theory: geometric quantisation, constraint quantisation and path integral quantisation. For a non-standard gauge theory things are more complicated, and the relational quantisation technique of Chapter 9 is taken to constitute, as yet, the only viable methodology.

We have already considered the problem of quantising a gauge theory within non-standard structure in some detail and a particularly important conclusion has been that the problem of providing a conceptual basis for any quantisation procedure seems to be intertwined with questions of interpretation related to the classical theory considered on its own. Most significant has been the relationship between reduction and quantisation procedures when applied to canonical gauge theories within which the Hamiltonian is itself a constraint. Problems in the interpretation of the reduced phase space in the context of a non-trivial representation of dynamics were connected to arguments towards the non-applicability of standard quantisation techniques. In addition to this line of argument from the conceptual analysis of classical theory to the formal structure of quantisation techniques, there is also significant scope for the converse; arguments from the formal structure of quantisation techniques to the conceptual analysis of classical theory. In particular, in a standard gauge theory at least, one may argue towards the primacy of the reduced phase space on the basis of the predication, to varying degrees, of all three quantisation techniques upon its symplectic structure. It is this second style of argument that could perhaps be seen to lend itself particularly well to the support of a structural realist stance with regard to the ontology of physical theory.

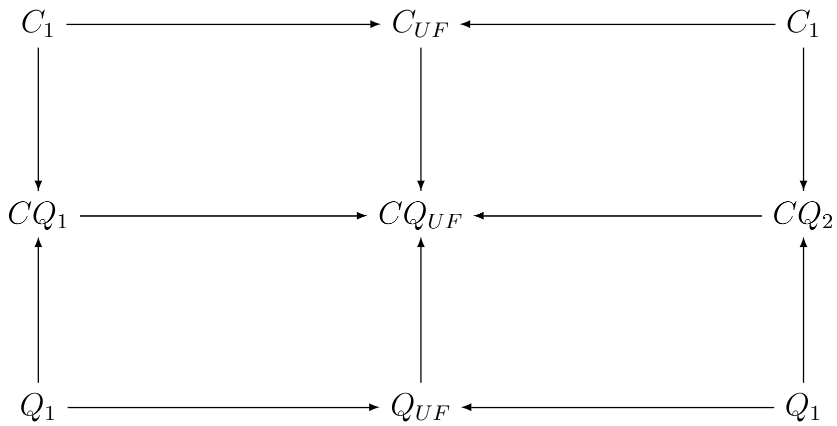
Quantisation is a bridge between classical and quantum theories and thus provides a direct and rigorous way of linking historically successive theories.[78] That the bridge itself is found to point to certain structures within the predecessor theory as in some way essential to that theory is extremely interesting. Briefly restating one of the principle motivating arguments of ontic structural realism (OSR), it is observed that throughout the history of science empirically successful theories are often, if not always, replaced by theories which include starkly different types of theoretical entities and objects. From this we may make the pessimistic, meta-inductive leap to the conclusion that the terms included within our current best theories that relate to theoretical entities and objects should not be thought of as constituting a genuine, robustly referential ontological vocabulary. Rather, the proponents of OSR contend, we should focus our attention on the structural aspects of physical theory and attempt to reconceive the notion of what constitutes the ontological vocabulary in terms of the structure common between successive theories. If the formal structure of quantisation techniques itself points to certain key structural facets of classical theory then it seems natural to ask what these structures correspond to within the quantum theory. We may then be able to specify precisely the structures that, according to OSR, should be reified when constituting a structural scientific ontology at the classical/quantum boundary such that it is robust to the challenge of pessimistic meta-induction.

A further motivation for OSR that has been considered in some detail within our discussion above was that based upon formalism underdetermination. Again restating some

[78] An alternative structural approach to conceiving of an ontology at the classical/quantum boundary would be to focus upon the classical limit of the relevant quantum theories. We will here neglect a detailed consideration of this option since it would provide little insight into the inter-formulation issue which we wish to investigate in parallel.

key ideas from above, we can understand the *interpretation* of a theory as pertaining to the demarcation of representational capacity within the theories theoretical structure (both observational and non-observational aspects). The *formulation* of a theory then relates to the manner which the theoretical structure of a theory has been expressed. We also have with distinct formulations, as opposed to distinct interpretations, that they are necessarily accompanied by a rigorous translation dictionary which allows us to transform from the language of one formalism to the language of the other. Formalism underdetermination is then the situation whereby two or more formulations of a theory exist such that they are either more or less restrictive as to which interpretations can be applied. If a traditional realist notion of scientific ontology is then appealed to, we are led to a situation of ontological underdetermination which is specifically driven by diversity with the formulations of a theory. The putative resolution of this problem from OSR is to reconceive our notion of ontology such that it is now constituted by structure common between two formulations. To be defensible as an ontology, however, it seems reasonable to insist that this common structure should take the form of a unifying framework rather than merely an interrelation. Here again it is possible that the formal machinery of quantisation may prove extremely important as a guide to identifying the right structure.

Let us assume we are given two formulations of a classical theory which have been quantised (perhaps by different methodologies). We would presume that from the two classical formulations will result the same quantum theory (although this is not guaranteed) and we would thus then have two quantum formulations of this single theory. Let us denote these formulations as C_1, C_2, Q_1, Q_2. A genuine implementation of the OSR programme for resolving cases of underdetermination would then provide us with a unifying framework for each of the pair of formulations, C_{UF} and Q_{UF}. Furthermore, a genuine implementation of the OSR programme for confronting the challenge of pessimistic meta-induction would give us a structural bridge between each of the classical and quantum formulations: CQ_1, CQ_2. It should also give us such a bridge between our two classical and quantum frameworks: CQ_{UF}. And furthermore, these two unifications should cohere. We can illustrate the situation graphically (committing a small abuse of mathematical diagrammatic convention) as follows:



Implementation of such a complex schema might be assumed to be impractical in general terms. However, armed with the mathematically well-defined quantisation procedures and interrelations between formulations we may perhaps be able to make some progress. In particular, it is highly suggestive that the symplectic and observables structure common between Lagrangian and Hamiltonian formulations at a classical level, is paralleled by the inner product and observables structures which are common between formulations at a quantum level. In the next section we will reconsider these issues in both more detailed and more concrete terms.

Chapter 17

Case study I: Lagrangian and Hamiltonian formalisms

17.1 What is underdetermined?

Let us consider a classical system consisting of a finite number of degrees of freedom and assume that this system does not display any local symmetry.[79] The physical theory describing such a system is Newtonian mechanics and in modern terms the two principal formulations available are Lagrangian and Hamiltonian (unfortunately, we do not here have space to consider the Hamilton-Jacobi formulation also). The Lagrangian formulation of Newtonian mechanics is framed within the space of solutions to the Euler-Lagrange equations which are dynamical curves, $\gamma_{EL} : T\mathcal{C} \to \mathbb{R}$ in the velocity-configuration space (the tangent bundle), $T\mathcal{C}$. The Hamiltonian formulation of Newtonian mechanics is framed within phase space (i.e., the cotangent bundle $\Gamma \equiv T^{\star}\mathcal{C}$) with Hamilton's equations picking out a preferred tangent vector field on phase space, X_H, which is sufficient to define the set of dynamical curves for any specification of instantaneous initial data.

By the criteria and definitions detailed in §15.1 and §15.2 what we are dealing with here is two distinct formulations of the Newtonian theory of mechanics: neither Lagrangian nor Hamiltonian formalism furnishes us with an ontology without a further interpretation and the two are connected by a rigorous translation dictionary provided by the Legendre transformation together with the set of maps (parameterised by a one dimensional time parameter) that exists between a given solution in the Lagrangian formulation and the corresponding sequence of instantaneous states in the Hamiltonian formulation. The crucial question, in light of our above analysis, is then whether we should understand these formulations as leading to an underdetermination of the relevant ontology. This de-

[79] Here and below we neglect the role of global symmetries for the sake of brevity. An analysis of the structural connections relevant to them would follow straightforwardly from what we say about observables and state spaces.

pends on the nature of the relevant interpretations available and their relationship to these two formulations.

Focusing in particular on the temporal ontology of Newtonian mechanics, two candidate interpretations are available. The first is constituted by the classic Newtonian picture of instantaneous states of the world together with deterministic laws sufficient to fix all past and future states given an *initial* state. We will call this the *instantaneous* picture of the world and understand it as specifying an ontology which posits instantaneous states as part of the fundamental *furniture* of the world. Supplementary to this picture we can ascribe additional and *more metaphysical* structure such as a dynamical notion of time and an ontologically privileged present (Markosian (2011)). Our concern here is not with the detailed philosophical analysis of these additional interpretational structures and the extent to which they prove acceptable additions to the project of furnishing the relevant theory with an ontology. They are certainly not generally taken to be precluded by Newtonian mechanics at least.[80] Rather, what we shall assume to be at the very least non-controversial is that given the viability of an interpretation in terms of a instantaneous picture, one may – if it is deemed reasonable – supplement this interpretation with additional temporal ontological structure such as a dynamic time.

A second interpretation of Newtonian mechanics that provides us with a distinct temporal structure is in terms of entire four dimensional histories which are specified by *atemporal* laws (i.e., laws that are not defined at a given time) together with initial and final boundary conditions. We will call this the *teleological* picture of the world since it implies the final boundary data is fundamental in determining the laws. Unlike the instantaneous picture is does not necessarily posit instantaneous states as part of the fundamental furniture of the world and, relatedly, it is not necessarily as amenable to supplementation with the additional *more metaphysical* structure mentioned above and discussed in more detail below. We do not mean this as a particularly strong claim and will not therefore seek to make a justification of it in a strong sense. Rather, we believe it is at the very least reasonable to assume that an interpretation of Newtonian mechanics in terms of a teleological picture is at face value going to look more like the non-dynamic, eternalist type stance as to the metaphysics of time and less like the dynamic/privileged present type stances.

An illustrative example of a potential association between the two pictures and a de-

[80] See Wüthrich (2010) for an interesting discussion of the extent to which the presentist view *is* precluded by theories of quantum gravity of exactly the type that have been extensively detailed in this work.

bate in contemporary *strong metaphysics*[81] is provided by the case of presentism already mentioned above. We can summarise basic relationship between our two pictures and the two sides of the presentism debate using the following tables[82]:

instantaneous picture	**Teleological picture**
Instantaneous states are fundamental	Spacetime is fundamental
Dynamical laws are defined for a given temporal state	Dynamical laws are defined atemporally
Initial conditions are fundamental	Initial and final conditions are fundamental

Presentist Stance	**Anti-Presentist Stance**
Privileged present	No privileged present
Dynamism (i.e., real temporal flow)	No dynamism
Only the present exists	Entire space-time 'block' exists

Just as there is a clear intuitive relationship between the aspects essential to the instantaneous picture and the presentist stance, there is a clear intuitive relationship between the aspects essential to the teleological picture and the anti-presentist stance. It would seem, furthermore, that the teleological picture is such that it is inherently hostile to presentism – the laws, boundary conditions and fundamental objects are things that, by the presentist lights, do not exist. Thus, at a superficial level of analysis at least, there is a natural way of *cashing out* the difference between our two pictures in terms of a substantive metaphysical difference.[83]

Even if we were to be more minimalist as to the level of metaphysical structure we wish to permit, then we may still end up with genuine differences between the two pictures. Whereas the instantaneous picture is predicated upon an ontology that necessarily

[81]By this we mean metaphysics of variety whereby ontological assertions concerning objects and concepts that go beyond usual scientific discourse are treated in the *thick* sense of Ontology rather than just more deflated ontology.

[82]Thanks to Sam Baron for help with this.

[83]There is also reasonable scope to understand the difference between the instantaneous and teleological pictures as possibly grounding a fundamental metaphysical difference as to the laws of nature. For example, it has been claimed that the disposition essentialist viewpoint on laws of nature is inconsistent with the principle of least action that is fundamental to the teleological picture Katzav (2004). See (Smart, 2012, §8) for further discussion.

includes instantaneous states as fundamental, the teleological picture is not *necessarily* predicated upon such an ontology. Thus any approach to space-time ontology which *precludes* fundamental instantaneous states can only be reconciled with the teleological picture – and is thus more naturally at home within the Lagrangian formalism. Such an argument is of course not sufficient to establish that there is no reasonable conventional realist ontology that transcends the Lagrangian/Hamiltonian divide. Rather, we have that there are at least *some* notions of ontology that are underdetermined by the case in hand, and thus that there is a requirement for the proponent of OSR to provide a viable alternative ontology, even if there is not an acute problem for the realist in pointing to possible ontology that is *is not* underdetermined.

We are now in a good position to examine our first test case for the possibility of formalism underdetermination. We have two formulations of a theory together with two viable and distinct interpretations. Above we listed three ways in which such a situation may lead to formalism underdetermination. Firstly, the underdetermination may be *strict*, meaning that there is no single interpretation that can be applied to both formalisms. Secondly, it may *exclusive*, meaning that there exists at least one interpretation which is applicable to one formulation but not to another. Thirdly it may be *loose*; meaning that one or more of the interpretations are more natural to one formulation than to another.

Since the teleological interpretation (or interpretation type) is applicable to both Hamiltonian and Lagrangian formulations the first does not apply. One could argue for the second on the grounds that the instantaneous interpretation might seem not to be applicable to the Lagrangian formulation. However, one may reconstruct the Lagrangian formulation such that it is based upon points rather than curves within the tangent bundle and such that the dynamical equations are differential equations giving a unique specification of dynamics at such a point rather than restrictions on possible curves. Such a re-conception means that it is possible to apply a instantaneous interpretation to the Lagrangian formalism. However, the historically prior and arguably most fundamental way of understanding Lagrangian mechanics is in the context of action principles and variational calculus and such formal structure does necessarily lead to a formulation which is in terms of curves with two boundary conditions. This point will be further born out when we come to discuss the quantisation of Lagrangian mechanics in terms of path integral methods as well as the intimately related issue of symplectic structure. There is therefore a good case for the Lagrangian formulation being more naturally interpreted in teleological rather than instantaneous terms and thus for us being confronted with a loose case of formalism underdetermination.

Given that the solution space of the Lagrangian formulation is that of curves with two boundary conditions, the natural interpretation is one in terms of a histories based ontology; with the *furniture* of the universe entire four dimensional spacetimes along with the appropriate initial and final conditions (i.e., the teleological picture). Given that the solution space of the Hamiltonian formulation is an initial data space, the natural interpretation is in terms of a instantaneous state based ontology; with the *furniture* of the universe three dimensional spatial states with appropriate instantaneous data (i.e., the instantaneous picture). Since the two formulations are empirically equivalent and yet, to an extent, furnish us with distinct ontologies there is a challenge to the realist to break the underdetermination.

Since the underdetermination in question is only of the loose variety, one obvious realist response would be to question its legitimacy as a genuine case of underdetermination rather than break it; and there is perhaps a good case to be made on this score relying upon the unnatural but viable instantaneous interpretation of the Lagrangian formalism. However, since this is merely the first of three examples, and it will prove a useful heuristic for the strictly underdetermined third example, detailed discussion of such a realist counter may be justifiably set aside.

17.2 Scientific realist responses

As discussed above (§15.4), a number of strategies for underdetermination breaking are available to the realist, all of which amount to appeals to some form of external criteria. Following on from that discussion the two most viable criteria that a scientific realist would seem best advised to utilise are: i) an appeal to greater simplicity; and ii) an appeal to greater heuristic fruitfulness. The first of these was introduced in the context of work by North (2009) within which simplicity was understood specifically in terms of minimal mathematical structure. The case of Lagrangian and Hamiltonian mechanics is that which North discusses in some detail, so it will be useful to consider the specifics of her argument.

Reiterating from above, the basic premise of North's form of realism is that the minimal, geometrical, coordinate-free structure of a physical theory is what is real. From this she argues that whichever formulation of a theory utilises the minimal amount of such structure should be taken as the true one. As we have discussed extensively above the

Hamiltonian formalism rests on the presumption of a symplectic structure. What North points out is that this symplectic structure is merely associated with a volume element and the Lagrangian formalism, on the other hand, presumes a metric structure which gives a distance measure. Thus, since a metric structure presumes a volume structure but not vice versa, we can argue that the Hamiltonian formulation of mechanics contains the fundamentally minimal, symplectic structure. Therefore, if we accept that structural minimalism should be conceived of as a viable underdetermination breaking criteria, then we should take the Hamiltonian formulation as that which is associated (via the appropriate interpretation) with the true ontology of the world. This would lead us to a (admittedly weak) argument that the temporal ontology of the world conforms to the instantaneous rather than teleological picture.

Now, for this case in particular North's argument in favour of the Hamiltonian formulation of Newtonian mechanics does seem fairly convincing – and we shall below concur with her idea that symplectic structure is what is fundamental to understanding this first case. However, as was argued in general terms above there is a severe difficulty with any view which seeks to carve away any and all seemingly excess mathematical structure that is associated with the formal aspects of physical theory. By following North's prescription in our particular case we would be driven to relegate both the Lagrangian formulation of mechanical theory and the metric structure which it presumes to a non-fundamental level within our theory – and, given the predominance of metric structures in special and general relativity, if such a move had been taken seriously by Einstein then it would have been a massive impediment to the development of relativistic spacetime physics in the early twentieth century! A general precept to always dispense with non-minimal structure would seem to tie one hand of the creative scientists since such structure is always potentially a fruitful resource for future development. This leads us to the second prospective underdetermination breaking criteria that a scientific realist might hope to appeal to: heuristic fruitfulness.

To an extent this second criterion seems to support both North's choice of fundamental structure and *true* formalism. As should be familiar from above it is precisely Poisson bracket structure associated with the symplectic form within Hamiltonian mechanics that is central to geometric quantisation. And in the early days of quantum mechanics – particularly in Heinsberg's formulation – this same structure was also key. Thus, ignoring relativity theory, one may be able to argue that for the first of our test cases the two best criteria for privileging one formulation over another cohere: Hamiltonian mechanics is both simpler and proved more fruitful. However, it is difficult if not impossible to see how one may neglect the pivotal relationship between Lagrangian theory and Einstein's work of relativity and as we have already stated, this argument for Hamiltonian theory in terms of the heuristic value of the Poisson bracket can be counter weighted by an argument for the heuristic value of the metric structure of Lagrangian theory. Moreover, the Lagrangian theory was eventually also quantised via path integral methodology and thus its structure has proved fruitful within both branches of modern physics. What should be taken as the abiding lesson from the consideration of the relative heuristic value of Lagrangian and Hamiltonian formulations of Newtonian mechanics is not that one or the other should be privileged – but rather that there is invaluable utility in diversity. It is the large scale heuristic plasticity that is enabled by a bipartite system of mechanics that can perhaps at least partially be said to account for the huge expansion of physics in the early twentieth century. Thus, rather than giving us good reason to break the formalism underdetermination, heuristic considerations, in this case at least, seem to impel us to embrace it. Arguments based on heuristic fruitfulness in this case give us reason to reject scientific realism in favour of preserving the ontological underdetermination with which it is inconsistent.

17.3 Quantisation and the structuralist response

The ontic structural realist response to cases of formalism underdetermination is to seek to reconceive the relevant notion of ontology in structural terms such that it is no longer underdetermined. As discussed above (§15.5), for such structure to genuinely constitute an ontology it is required to consist of more than a mere interrelation between formulations, we need to find a suitably generalising physico-mathematical framework which includes the requisite level of dynamical structure. Is this possible for the case of Lagrangian and Hamiltonian mechanics? Based on the analysis of Belot (2007), we can make a good argument that the answer is yes. As has already been partially discussed above (§5.4,§11.3), Belot's work illustrates that for standard theories of mechanics (i.e., standard gauge theories and non-gauge theories): the space which represents unique solutions within a Lagrangian formulation of mechanics, has a close formal relationship with the space which uniquely represents instantaneous states within a Hamiltonian formulation. Within Newtonian mechanics these two spaces are simply the space of curves solving the Euler-Lagrange equation, $\gamma_{EL} \in \mathcal{S}$, and phase space, $T^{\star}\mathcal{C}$. Not only are these two spaces connected by a set of maps between time slices of Lagrangian solutions and instantaneous canonical states, but since $\mathcal{S}$ is, like phase space, a symplectic manifold, it is possible to prove that the two relevant dynamical arenas are *symplectically isomorphic*. The existence of this symplectic isomorphism then allows us to fix a precise relationship both between functions representing observable quantities within the two formalisms and between the representation of dynamical change in the observable quantities. Given a preferred slicing of a Lagrangian solution, for every moment of time we can construct a symplectic isomorphism between a phase space function and a corresponding function on $\mathcal{S}$ – and this relationship allows us to understand both functions as representing the same underlying physical quantity as it changes over a dynamical history. Thus the mutual symplectic structure of Lagrangian and Hamiltonian mechanics provides us with exactly the kind of generalising framework, including dynamical structure, which we are looking for and although we will certainly not claim that this analysis is complete,[84] there is a convincing case for an ontic structural realist account of the Hamiltonian and Lagrangian formulations of Newtonian mechanics in symplectic terms. This ontology is not constituted by the symplectic isomorphism itself but by the interconnections between dynamical structures that it encodes at the level of both observables and the state spaces. To accept this ontology is not to endorse either the instantaneous or teleological interpretations, rather through OSR we are able conceive of a fundamental reality that stands behind these two contrasting pictures of the world in terms of precise structural framework.

We then come to the question most crucial to our analysis. Is this prospective structural ontology of the suitable type to deal with *both* underdetermination issues and the historical undermining of pessimistic meta-induction? Would it be appropriate to conceive of the relevant symplectic structure as preserved between classical and quantum mechanical arenas? Again, our question is to an extent already answered. In our discussion of geometric quantisation techniques (§3.1) it was noted that one of the key steps was defining the map $A : f \to A_f$ which takes us between classical algebra of observables, defined by functions on a symplectic manifold, and the quantum algebra of observables, defined by self adjoint operators on a Hilbert space. One of the restrictions on this map was that $[A_f, A_g] = i\hbar A_{\{f,g\}}$ and thus we see that by definition the geometric quantisation scheme is such that the classical Poisson bracket structure is carried over into the quantum

[84]An interesting and important extension to our analysis which would strengthen our case, would be a full illustration of how the relevant generalising framework might be constructed in precise semantic terms. We leave this task as well as the corresponding analysis for the other case studies to future work.

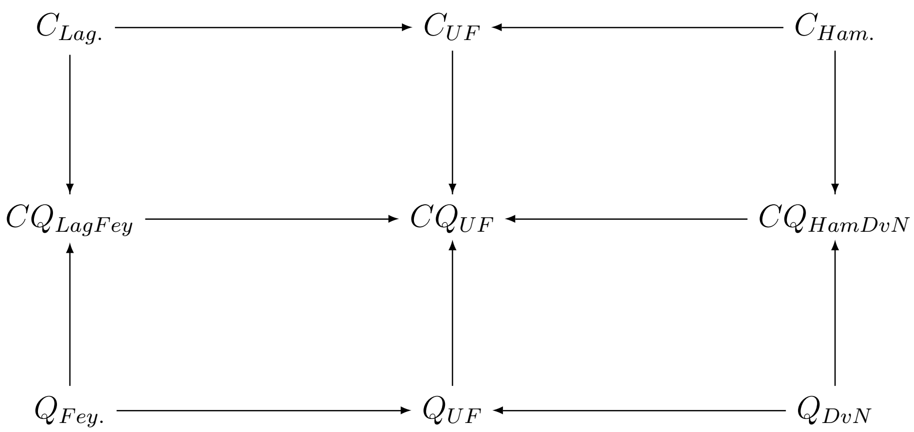
context in terms of the commutator. We can therefore justifiably argue that there exists a structural bridge between the observables of classical Newtonian mechanics and non-relativistic quantum mechanics at a formal level, precisely in terms of the link between the binary operations constituted by the Poisson bracket and the commutator. This analogy is also reflected at the level of dynamics since when combined with the Hamiltonian observable it is the binary operation that is responsible for generating evolution. Further to this structural bridge from the symplectic form to the commutator, there is also a suggestive analogy between the classical state space (a manifold equipped with a symplectic structure) and the quantum state space (a vector space equipped with an inner product structure).[85]

Independently of anything to do with formalism underdetermination, a proponent of OSR would therefore argue that the fundamental structure of a classical or quantum theory is related to maps between algebras of observables, the relevant binary operations and the relationship between the classical and quantum state spaces. Fundamentally this is what is structurally consistent between the classical and quantum theories. It is therefore what OSR implies we should seek to reify in the face of pessimistic meta-inductive arguments. However, this is also, roughly speaking at least, the type of structure which we were driven towards when considering the ontology of the classical theory alone so there would seem to be *prima facie* coherence in our approach.

Let us then examine the case in hand more carefully. Given our two classical formulations we arrived at a 'structural ontology' encoded by a symplectic isomorphism between both the relevant observables and state spaces. Given a generalised, geometric picture of classical and quantum theory we arrived at a structural ontology encoded by: i) a Lie algebra morphism (up to a factor) connecting the algebra of observables and the relevant binary operation; and ii) the connection between the symplectic and inner product structures. Although these are not *the same* structures, they are closely related. One way to refine our analysis a little is to consider two different formulations of quantum theory, look at the common structure, and compare this to both the classical-classical formulation common structure and the general classical-quantum common structure. If we presume to have quantised the Lagrangian formulation classical mechanics using a path integral methodology and the Hamiltonian formulation using canonical quantisation (which just amounts to a concrete implementation of geometric quantisation) then we would have two

[85]This connection is undoubtably a subtle one and we do not have space here to consider it in full formal detail. See future work for a detailed analysis of this aspect to our scheme.

formulations of quantum theory, each based on a formulation of classical theory. We will label these two formulations after their principle originators – Feynman on the one hand and Dirac-von Neumann on the other. Our desired consistent structural ontology could then expressed using the diagram that was introduced above:



In this notation, our discussion thus far has already effectively covered C_{UF} and CQ_{HamDvN}. We will now briefly consider the rest of the diagram in order to give at least a superficial evaluation of the extent to which the relevant structural notion of ontology is suitably 'commutative'. The fundamental dynamical equation within Feynman path integral quantum mechanics is, for a single particle:

$$Z = \langle q_f | e^{-\frac{i}{\hbar}H(t_f - t_i)} | q_i \rangle = \int Dx e^{\frac{i}{\hbar}\int_0^T \mathcal{L}(q,\dot{q})dt} \tag{17.1}$$

Where D is the functional measure. This path integral expression describes quantum mechanical behaviour in a configuration space in that, roughly speaking, it gives us a probabilistic weighting to paths through that space between an initial position q_i and a final position q_i. We thus see that, under the Feynman approach, a quantum system is associated with a space of possible histories (i.e., the space over which the integral is taken) and the nature of the path integral is such that it gives (in an informal sense) an inner product structure to that space.

Within the classical theory we also focused upon a space of histories as fundamental to the Lagrangian formulation; and it was the symplectic structure of that space which we took to constitute one side of the structural bridge between Lagrangian and Hamiltonian theory. Furthermore, in the generalised abstract case and the case of Hamiltonian theory, there is an extent to which the symplectic structure within the classical theory is analogous to the inner product structure within the relevant Hilbert space. It is natural therefore to ask whether the symplectic structure of the classical history space in Lagrangian theory can be connected with a Hilbert space, together with the necessary inner product structure, within Feynman path integral quantum mechanics.

Unfortunately, although its heuristic, intuitive and practical value is undoubtably great, the Feynman path integral as it has been introduced, is insufficiently mathematically well-defined for us to be able to answer this question. Consideration of the project of providing a more rigorous mathematical basis to it would take us far beyond the limits of our current discussion, but we may at least note that according to Albeverio *et al.* (2008) the Feynman path integral for the solution the Schrödinger equation can be interpreted rigorously as a *Fresnel integral*[86] over a Hilbert space of continuous paths. Thus, given a suitable formalisation, it does appear to be correct to think of path integral quantum mechanics in terms some form of *Hilbert space for histories*. There is, therefore, some formal support for a tentative proposal that a structural bridge may be made between Lagrangian classical mechanics and path integral quantum mechanics in terms of a connection between: a classical space of histories with symplectic structure, on the one hand; and a quantum space of histories with an inner product structure, on the other. We do not, however, have the Lie algebra morphism that can be demonstrated to connect the observables and dynamics of the classical Hamiltonian theory with the Dirac-von Neumann quantum theory (as arrived at via canonical quantisation). Relating the classical Lagrangian notion of observable to some precisely analogous structure within path integral quantum theory – if it is possible – is a highly non-trivial challenge.

In addition to seeking this structural connection between classical Lagrangian and quantum path integral formalisms, consistency with the OSR philosophical framework drives us to look for a similar connection between path integral and Dirac-von Neumann quantum formalisms. Not least this is because these two quantum formalisms would appear to be naturally associated with interpretation in terms of disparate ontologies – a quantum teleological type and quantum instantaneous type picture respectively. Further to this, in order to establish the relevant commutativity we need to find a quantum unifying framework to parallel our classical unifying framework and then hope that the structural commonalities between these two frameworks (the middle edge of our diagram) mirror those between the individual classical and quantum formulations (the two outside edges).

[86] A special type of oscillating integral defined on a real vector space equipped with a norm.

Unfortunately, our progress is once more hampered by the unsolid mathematical basis of Feynman's approach. Again what is desired would be a well-defined Hilbert space of histories which could then be connected to the traditional Hilbert space of instantaneous quantum states. In such circumstances, if the two Hilbert spaces could be shown to be unitary isomorphic *and* the relevant isomorphism can be understood as entwining the representations of two sets of quantum observables, then we would have established, despite the apparently fundamental interpretational difference, that the two quantum formulations are fundamentally manifestations of the same underlying physico-mathematical framework/structure. The situation with regard to the Hilbert space aspect of our problem is again promising. According to Dowker *et al.* (2010) we may formalise a histories approach to quantum theory using the framework of quantum measure theory (Sorkin (1994)) and proceed to construct a histories Hilbert space which can be proved (given a unitary quantum theory with a pure initial state) to be isomorphic to the conventional Hilbert space of the Dirac-von Neumann formalism. However, despite this success at the level of state-spaces, the situation with regard to observables is less promising as there is currently not a sufficiently general procedure for constructing an observables algebra within a histories Hilbert space formalism, let alone a proof that such histories observables are suitably related to their conventional Dirac-von Neumann counterparts.

We are, therefore, not in a position to reach a strong conclusion with regard to the observables aspect of a cross-formulation quantum mechanical structural framework – and according to our own criteria this means we have not quite met the necessary conditions for an adequate structural ontology at the quantum level. However, through the relevant state space connections we have suggestive evidence that our application of OSR in terms of the digram above is leading us in a promising direction. In particular, for all of the four outer nodes of the diagram – i.e., the Lagrangian and Hamiltonian formulations of classical mechanics and the path integral and Dirac-von Neumann formulations of quantum mechanics – all the necessary structural connections can be seen to hold with regard to the state spaces involved. The symplectic structure and Poisson bracket algebra of observables are what is fundamental at a classical level, the inner product structure and commutator algebra of observables are what is fundamental at a quantum level. The classical and quantum structures are analogous in the case of the state spaces and, modulo the difficulties mention, connected directly by a Lie algebra morphism in the case of the observables. More work must be done to further refine details of this project, but at this initial level of analysis at least, it seems we have good evidence for the fundamental consistency within our OSR-style reading of Newtonian mechanics. Let us press on to our second case study to see if such success is replicated within standard gauge theories.

Chapter 18

Case study II: reduced and unreduced formalisms

18.1 What is underdetermined?

We now return our analysis to the standard gauge theories considered at great length in Part I. As should be familiar from that discussion, these systems are represented canonically through the constrained Hamiltonian formalism and are distinguished by featuring first class constraints that *can* be understood as purely generating gauge symmetries. We can understand the collection of these constraints as defining a sub-manifold, (Σ, ω), within the 'extended' phase space, (Γ, Ω), called the constraint manifold or physical phase space. The physical phase space is such that equivalence classes of points can be defined (via the null vector fields associated with the constraints) and we call these equivalence classes gauge orbits. In standard gauge theories it is physically reasonable to classify points that lie along a gauge orbit as physically equivalent in that they represent the same instantaneous state of the world. This in turn justifies the viability of switching to a reduced formalism, where through a quotienting procedure we construct a reduced phase space, (Π_R, Ω_R), with every point on a gauge orbit mapped down to a single point. The map between the physical and reduced phase spaces, π, is such that a version of both the Hamiltonian function and the Poisson bracket algebra of observables is carried over onto the reduced space. We thus have available two formulations of the mechanics of a standard gauge theory – the unreduced and reduced. They are empirically equivalent and connected by a rigorous translation dictionary as defined via the map π.

The crucial question is then, do these different formalisms precipitate a case of ontological underdetermination? To find out the answer we must consider the relevant interpretational structures that can be added to our bare formalism. This depends upon the theory at hand. Given that we have already laid much of the ground work in part three, it will best behove us to consider the case of the momentum constraints of general relativity

– an analogue of what we will say should be applicable to any standard gauge theory with a suitable change in the objects playing the role of *background structure*. Recall from above that these constraints can be associated with the action of the three dimensional diffeomorphisms group on a three dimensional spatial geometry. A point within the relevant reduced phase space (super-phase-space) can then be understood as representing an equivalence class of such three geometers or equivalently a single, diffeomorphically invariant object.

We can define within the unreduced formalism a number of interpretational stances as constituted by a position with regard to the ontological status of spatial points. One may be a spatial substantivalist and assert that such points are fundamental; or a spatial relationalist and deny this fundamentality. Whereas, within the unreduced formalism substantivalism may be achieved with or without an additional commitment to Haecceitism – one may be a straightforward or sophisticated substantivalist; within the reduced formalism, the two available substantivalist stance is the sophisticated Haecceitist. Thus, we see that which formalisms you use places restrictions on which interpretation (and therefore which ontology) is available.

If one wishes to ascribe stronger metaphysics to one's physics then one may make this underdetermination even more acute by cashing out the excess possibilities available only within the unreduced formalism in terms of some form of modal realism – formally this is because the two spaces contain different *cardinalities* of possibilities. Thus we see that in general there is considerable grounds for believing the case of reduced and unreduced formulations of a (standard) canonical gauge theory constitute an example of formulation underdetermination according to the terminology introduced in the last chapter. For the traditional doctrine of scientific realism to be applicable either this this underdetermination needs to be broken or an interpretations that can be applied to both formalisms must be accepted. As with the first case study, we will first consider the second option, and then move on to consider the structural realist response. The first option we will simply admit as a consistent, alternative to the picture presented here.

18.2 Scientific realist responses

Let us then consider the viability of the various arguments available to the scientific realist who wishes to privilege one of the two formalisms over the other. First, let us briefly reconsider the idea of making an appeal to metaphysics since one of our general worries about these strategies is well illustrated by the case in hand. The most obvious metaphysical resource that could be employed to break our second case of underdetermination would be some principle that compels us to: either reify the maximum or minimum possibility space; or endorse the maximal or minimal amount of absolute structure. Only such metaphysical principles would seem relevant and powerful enough to break the underdetermination. However, to endorse any such principle as the arbiter for our case would to enter into a vicious circularity – the choice between more or less possibilities/background structure is exactly what is underdetermined! Thus, it seems metaphysical principles can be of no comfort to the realist in constructing a rational basis for privileging either formalism.

Going from the more metaphysical to the more pragmatic end of the philosophical spectrum, the interpretational flexibility of the unreduced phase space could be taken as practical ground for preferring working with that formalism since it is the 'neutral base' (Rickles (2008)) from which to work. However, such pragmatic arguments are insufficiently strong to justify the type of *thick* realism that the scientific realist (by our definition) requires. They need to give an argument why one formalism rather than the other leads to *the true picture of the world* and pointing to the utility of working with whichever choice allows us to best hedge our metaphysical bets is clearly not enough to do this. Furthermore, as was argued in §15.4, flexibility is a double edged sword since (at least under some accounts) the more flexible the scientific framework is, the less explanatory value it can be understood as having.

In our general discussion we singled out simplicity and heuristic fruitfulness as the most viable science based principles for underdetermination breaking. We can think of the former principle precisely in the context of the North (2009) formulation that, as discussed above, is based on the idea of minimal geometric structure. Dynamics within the unreduced formalism requires for its definition the quadruple $(\Gamma, \phi_i, \Omega, H)$ with ϕ_i the constraint and H the Hamiltonian. Based on these four objects we can then either define observables and evolution via a weak commutation relation or upon the physical phase. Since in the unreduced formalism the constraints have been eliminated it can be defined simply via the usual triple of an unconstrained system – i.e., in this case (Π_R, Ω_R, H_R). We can thus give a precise sense in which the reduced formalism is simpler than the unreduced formalism. However, as has been noted before, the endorsement of such minimal structure arguments seems to be contrary to the history and practice of physics. The supposedly surplus structure of physical theory has proved, and is therefore likely to continue to prove, essential to theoretical development. This is particularly true when considering the surplus structure of the unreduced canonical formalism – it is precisely this surplus structure that is pivotal in Dirac route towards quantisation (§3.2). Moreover, it is in general extremely difficult to construct explicitly a true reduced phase space (i.e., a manifold with symplectic structure) without running into serious formal issues: superspace, for example, has problems with non-trivial topology. Thus in practice the quantisation of a standard gauge theory nearly always proceeds via utilisation of exactly the excess structure (i.e., the constraints) that simplicity arguments would lead us into dispensing with.

This last point might indicate that we may be able to make an argument for the unreduced space formalism upon heuristic fruitfulness. Such an argument is especially pertinent to our discussion given the indispensability of the unreduced formalism for the quantisation of non-standard gauge theories as detailed in Part II and Part III. However, for standard gauge theories at least, the symplectic structure of the reduced space is also an important heuristic tool for quantisation. In particular, it is precisely our ability to think of Faddeev-Popov quantisation of a standard gauge theory as a Feynman path integral on the reduced space that provides the conceptual basis for that technique. Thus, both spaces have proved to be able to provide us with heuristically useful structures and therefore neither can be privileged on the grounds of heuristic fruitfulness. In fact, the structure that seems most appealing from a heuristic for quantisation perspective is the connection *between* the spaces. This leads us naturally to consider the potential application of the ontic structural realist programme along the lines discussed in §15.6 and §16.3 above.

18.3 Quantisation and structuralism

Given the reduced and unreduced formulations of a standard gauge theory we would first like to consider the relevant structural connections at a purely classical level. Again what we are looking for is more than merely an interrelation between the formulations, rather we need to find a suitably generalising physico-mathematical framework including dynamical structure. This can be done fairly easily for the case in hand. The map π not only defines the relationship between the reduced phase space and the physical phase space, but is also necessarily such that it fixes a relationship between the relevant notions of observables *and* Poisson bracket structures. Consider the observable functions $f_R, g_R \in C^\infty(\Pi_R)$ and the Poisson bracket defined by the relevant symplectic structure

$\{f_R, g_R\} = \Omega_R(X_{f_R}, X_{g_R})$. We can then use π to pullback to an equivalent set of functions in the unreduced formalism, $f, g \in C^\infty(\Sigma \subset \Gamma)$ such that $f = \pi^* f_R$ and $g = \pi^* g_R$. Now, we have that π does not allow us to pullback the Poisson bracket structure of the reduced space uniquely, since $\omega = \pi^*\Omega_R$ gives us only a presymplectic structure. However, since the f, g are by definition such that they weakly commute with the constraints, the Poisson bracket associated with the symplectic structure of Γ will, when restricted to Σ, equip this collection of functions with the binary operation necessary for us to establish a symplectic isomorphism between the algebras. Explicitly (again following Faddeev (1969)) and assuming that Π_R is parameterised by coordinates Q and P:

$$\{f, g\}|_\Sigma = (\frac{\partial f_R}{\partial P}\frac{\partial g_R}{\partial Q} - \frac{\partial f_R}{\partial Q}\frac{\partial g_R}{\partial P}) \tag{18.1}$$

where we have suppressed the subtlety in indices needed for the lower dimension of the reduced space to be accounted for. Further to this first aspect of the classical unifying framework as constituted by the symplectic isomorphism between the algebras of observables it also trivially follows that the dynamics of the two formalisms can be suitably connected, and generalised. Since the relevant Hamiltonians are simply functions on either Π_R or Γ they are connected via π. Our dynamical framework is then just encoded in the equation:

$$\dot{f} = \dot{f}_R = \{f, H\}|_\Sigma = (\frac{\partial f_R}{\partial P}\frac{\partial H_R}{\partial Q} - \frac{\partial f_R}{\partial Q}\frac{\partial H_R}{\partial P}) \tag{18.2}$$

We have therefore established for the case of the unreduced and reduced formalisms, exactly the type of generalising physico-mathematical framework that the ontic structural realist would wish to reify. Satisfyingly this framework is of a very similar type as that discussed for the case of Lagrangian and Hamiltonian mechanics above. Although in this case we do not have a symplectic isomorphism between the relevant state spaces we do have such a relationship between the algebra of observables and the dynamical structure.

The next step of in our programme is to consider the quantisation of the two classical formalisms and then investigate the connections that exist at both classical-quantum and quantum-quantum level. Recall from above, one of the principal motivations for this exercise is to examine whether the structures that are common between two formulations of a theory are related to the structures that are common between predecessor and successor theories. Furthermore, if the ontic structural realist prescription for identifying ontology within physical theory is a consistent one, then we would expect a degree of commutation – the quantum-quantum structure should reflect the classical-classical structure we have just described. Assuming that we proceed to quantise the unreduced formalism via some variant of the Dirac methodology and the reduced formalism via a geometric quantisation technique then our structural commutation diagram should look as follows:

$$
\begin{array}{ccccc}
C_{UnRed.} & \longrightarrow & C_{UF} & \longleftarrow & C_{Red.} \\
\downarrow & & \downarrow & & \downarrow \\
CQ_{UnRedConst} & \longrightarrow & CQ_{UF} & \longleftarrow & CQ_{RedGeo} \\
\uparrow & & \uparrow & & \uparrow \\
Q_{Const.} & \longrightarrow & Q_{UF} & \longleftarrow & Q_{Geo.}
\end{array}
$$

Much of this diagram has, in fact, already been dealt with in our discussion. The exterior right hand CQ_{RedGeo} edge is merely the relationship between a non-gauge Hamiltonian theory and its quantum equivalent. From our discussion in §16.3 we have that the relevant structural ontology should be thought of as being encoded by: i) a Lie algebra morphism (up to a factor) which connects both the algebra of observables and dynamics; and ii) the (as yet not full explored) connection between the symplectic and inner product structures that defined the relevant states spaces.

The lower Q_{UF} is precisely the relationship between quantum formalisms reached via the Dirac constraint quantisation and reduced quantisation routes. From §4.1 we have that if the standard gauge theory in question falls inside the scope of a Guillemin-Sternberg conjecture proof then the two quantum formulations are related such that: i) the physical Hilbert space constructed through a constraint quantisation type approach, $\mathcal{H}_{phys}$, is unitarily isomorphic to that, $\mathcal{H}_R$, achieved by quantising the reduced phase space; and ii) the two quantisation procedures result in an equivalent set of observables to the extent that the isomorphism in i) also entwines the representations of the two sets of quantum observables (both of which can be connected back to the same set of gauge invariant classical observables).

Making the assumption that we *do* fall inside such a proof (this notably excludes cases such as the momentum constraints of general relativity and, without further simplifying assumptions, Yang-Mills theory) then we can give a clear characterisation of Q_{UF} in terms of the unitary isomorphism that encodes the relationship between both the states spaces, the observables and the dynamics. We can then in turn consider the relationship between this quantum unifying framework and its classical analogue (the middle vertical edge of the diagram). Although classically the relationship between the relevant state spaces is a little more subtle, the relationship between the observables – both quantum mechanically and classically – is represented via a map that encodes how the binary operations of each algebra are related. Furthermore, in each of the four cases (the four corner nodes) we have analogous inner product or symplectic structure essential to the definition of the state space. Thus, the essential structural commonality between the classical and quantum unifying frameworks must be understood in terms of: i) the maps encoding binary operations over algebras of obervables in each case; and ii) the relationship between the symplectic and inner product structures that define the state spaces . i) reflects precisely the structure identified more formally as a Lie algebra morphism when we were discussing CQ_{RedGeo} above and therefore evidences one aspect of the relevant $CQ_{UF} \leftarrow CQ_{RedGeo}$ link in the diagram.

The final two structural links that must be established for our diagram to *close* are those: a) between the unreduced classical formalism and the constraint quantised quantum formalism, $CQ_{UnRedConst}$; and b) between this structure and the general classical-quantum framework, CQ_{UF}, we have just identified. Starting, as is logically necessarily, with a) we must first decide whether it will be more valuable to our analysis to consider the informal version of constraint quantisation according to Dirac or its modern implementation in terms of RAQ or the MCP. Since we have already used the qualification that we are working inside the scope of a Guillemin-Sternberg conjecture proof it best behoves us to assume that the standard gauge theory we are considering is such that the Poisson bracket algebra closes with at worst structure constants. Thus, we can utilise RAQ for our analysis and make use of the group theoretic basis for both quantum and classical local symmetry which it provides us. Recall that classically the essential structures of the unreduced theory were taken to be given by the quadruple $(\Gamma, \phi_i, \Omega, H)$. What was not discussed explicitly above is that together the second and third of these also encode the structure of the constraint's Poisson bracket algebra since we have that: $\{\phi_i, \phi_j\} := \Omega(X_{\phi_i}, X_{\phi_j})$. It is the structure of this classical algebra that gives us a group theoretic basis for understanding the local symmetry in the theory and which gives the clearest structural bridge to the RAQ formalism. The quantum constraints within the RAQ scheme are taken to be represented as Hermitian operators acting on the auxiliary Hilbert space. The binary operation that defines the quantum constraint algebra is then defined by the commutator analogue of the Poisson bracket given by Ω. The quantum constraint algebra is then guaranteed to itself be a Lie algebra and we therefore have a Lie algebra morphism between the classical and quantum constraint algebras. Thus, the same type of structural connection that is essential to the relationship between the classical reduced and geometrically quantised formalism is key to the way in which the auxiliary Hilbert space is constructed in RAQ.

Furthermore, it is because the quantum constraints form a Lie algebra that that exponentiation of the constraint operators yields a unitary representation $U(g)$ of the corresponding Lie group G. We then have that the observables – including the Hamiltonian operator – are self-adjoint operators that commute with the action of this group (on the subspace Φ defined in §3.2.1). This is closely analogous to the definition of the classical observables in the unreduced formalism and means that the algebras of observables in each case have essentially the same structure. However, there is, of course, a second stage to the RAQ quantisation process – the construction of the physical Hilbert space. Significantly, there is no classical analogue in the unreduced formalism for this second step, in particular the inner product of $\mathcal{H}_{phys}$ has its correlate structure in the *reduced* classical formalism. It is therefore not entirely consistent to think of the quantum theory constructed via RAQ (or indeed any constraint quantisation methodology) as simply a quantum version of the classical unreduced formalism and to this extent, for this case our diagram is somewhat misleading. Nevertheless, to the extent to which this connection does make sense, its essential structure is encoded in terms of a Lie algebra morphism between the classical and quantum constraint algebras along with the associated relationship between the observables. Furthermore, to the extent to which it does not make sense due to the absence of a classical analogue for the inner product structure of the physical Hilbert space, we have resources within the reduced classical formalism that can fulfil the required role.

We thus have that the classical reduced and unreduced formalisms together might seem a better structural match for constraint quantisation. In this context, we may then be able to understand the RAQ rigging map η as something like the analogue structure to the classical reduction map π. However, as was discussed in §4.1 the connection between the two is more subtle that it might at first sight seem. More detailed investigation of this point is much warranted – in particular a careful formal examination of the structural connections between classical and quantum quotienting procedures would be very interesting. On a more informal level, clearly there is a sense in which what is achieved in RAQ is directly analogous to the quotienting of a Lie group, and this is precisely what we are understanding the classical symplectic reduction as achieving also. We could argue therefore that all the relevant connections between our classical and quantum formalisms are always encoded within: i) the relationship between Lie algebraic and group theoretic structures; and ii) the analogy between the symplectic and inner product structure of the state spaces. i) includes the observables, the Hamiltonian and the constraints themselves. Given the clear coherence of this picture, it would seem that this second case study leaves the ontic structural realist well placed to define exactly the required 'generalising physico-mathematical framework'. This framework essentially consists of the structure which encodes the relationship between key quadruple of state spaces, observables, dynamics, symmetries.

Chapter 19

Case study III: shape dynamics and complete observables

19.1 What is underdetermined?

We now return our discussion to the task of adjudicating between the two viable *denials* of time that were examined extensively in Part III. Recall that there are currently no empirical grounds for differentiation between the approaches towards canonical general relativity that were discussed and, furthermore, with regard to the treatment of momentum constraints there is also little or no inherent conceptual difference. Fundamentally, and at a formalism level, the choice as it was presented in its final form was between: i) a 3D conformally invariant theory with a fixed foliation and change generating global Hamiltonian constraint; and ii) a foliation invariant theory with the local Hamiltonian constraints replaced by a single master constraint and the observables constructed accordingly via application of the complete and partial observables Ansatz. Interpretationally, each of these formalisms was associated with a particularly interesting – and in many ways challenging – notion of ontology. The first formalism, which we shall refer to simply as shape dynamics, is naturally interpreted in terms of a Machian view whereby absolute structure with regard to space, time and length scale has been eliminated but a notion of change with respect to an internal, equitable measure of duration can be defined. The second formalism, which we shall simply call complete observables, is then naturally interpreted in terms of a (non-scale invariant) four dimensional non-local observables picture which not only eliminates absolute spacetime structure but also any notion of change whatsoever – we are left with a Parmenidian picture of reality.

The two formalisms are extremely closely tied to the two interpretations; the Machian view on time and scale in fundamentally inconsistent with the complete observables scheme and the Parmenidian picture of time is fundamentally inconsistent with both the preferred foliation and emergent time aspect of shape dynamics. This is not to say that in the case of either formulation there is *no* interpretational leeway – particularly with regard to space there is scope for alternative ontological pictures to be associated with each formalism. However, with regard to time there is an extent to which each formalism is *locked in* to a particular interpretation and it would certainly seem the case that *no single interpretation could possibly be applied to both formalisms*.

Furthermore, provided we subject our analysis to the very significant qualification that the complete observables scheme has been applied within the CMC foliable sub-set of solutions to the Einstein equation, we would expect a rigorous translation dictionary to exist between the two formalisms. We have this since the complete observables scheme is merely a prescription for defining observables within conventional canonical general relativity and (in the CMC gauge) this theory has been shown to be equivalent to shape dynamics. Thus, our third, final and most important case study can be reasonably understood as a strict case of formulation underdetermination. It could, of course, also be understood as a case of theoretical underdetermination since there is an important sense in which we may view shape dynamics as an alternative theory to general relativity, rather than an alternative formulation.

How seriously one views this issue depends to a large extent on the way one views the physical viability of non-CMC foliable spacetimes. If the non-physical nature of these solutions, or more precisely phenomena associated with such solutions, is accepted then our choice genuinely would be reduced to that between different formulations of the same theory and their associated interpretations. However, one could quite reasonably claim that such an assumption is not entirely warranted – and we would be better to treat shape dynamics as an alternative theory to general relativity with the choice between the two merely (weakly) empirically underdetermined. To someone who is insistent on this point there is perhaps no strong rebuttal available. Yet one may deflect the point of contention, however, by appeal to a comparable case.

As was discussed extensively in §10.2 covariant and canonical general relativity are, strictly speaking, empirically distinguishable since the former, and not the latter, is well-defined upon the non-globally hyperbolic subset of spacetimes. Thus, to the extent that it is problematic to call shape dynamics a reformulation of canonical general relativity, it is also problematic to call canonical general relativity a reformulation of covariant general relativity. Moreover, in moving to the shape dynamics formalism from the canonical formalism we are not strictly excluding all non-CMC foliable spacetimes but merely the globally hyperbolic, non-CMC foliable spacetimes. We can therefore see this issue may not be as serious as it may seem. Furthermore, given that, to the best of our knowledge, the universe that we live in is at large scales best modelled by the FRW-solution to the Einstein field equations – which *is* a CMC-foliable spacetime – and at very small scales not modelled well by general relativity at all, this entire question of the existing phenomena related to non-CMC classical solutions might be considered moot.

For the rest of our discussion we will simply *assume* CMC-foliabllity meaning that shape dynamics is an empirically equivalent formulation of canonical general relativity and therefore that we are dealing with formulation underdetermination. The crucial point is that this assumption does not imply that we have fixed a foliation for the complete observables scheme nor that the solutions we are dealing with cannot be invariant under refoliation. From the perspective of the complete observables formulation we are merely restricting ourselves to equivalence classes of solutions which have a CMC foliated representative. Thus we have that the duality we shall be exploring is between two formalisms with well-defined yet very different notion of symmetry. Given a spacetime (and ignoring 3D spatial diffeomorphisms) we have the option of understanding the relevant symmetry in terms of *either*: i) 3D conformal invariance up to global scale and reparameterisation invariance; *or* ii) refoliation invariance. These symmetries can only really be interpreted in terms of fundamentally different ontologies and thus precipitate an acute case of metaphysical underdetermination.

19.2 Realism: shape space or Parmenidian states?

The gamut of underdetermination breaking criteria available to the realist should now be familiar, as should also be their various weaknesses. For the case in hand, what might seem like the most obvious response would be some argument from physics enriched metaphysics that seeks to convince us that either formalism better implements: *background independence* (in the case of complete observables); or *conformal invariance* (in the case of shape dynamics). However, as has been noted several times, use of such arguments amounts to question begging; they are predicated upon a stance as to precisely the ontological difference that is underdetermined.

Arguably a better candidate is a more general prescription against background structure. To give scientific credence this could be conceived of in terms of a principle drawn from a reasonable amount of historical evidence as well as support amongst practising physicists. However, what counts as less background structure? We are comparing one approach with more of a temporal background, but almost no notion of absolute scale, to another with no temporal structure at all, but a well-defined notion of scale. To choose one or the other as more significant would seem arbitrary and might also be considered precisely the type of question begging we are trying to avoid.

Another alternative would be to invoke a minimal structure argument of the North (2009) type discussed extensively above. The most obvious candidate for such structure is the algebra of constraints since this is fundamental to each formalism and can be subjected to a precise notion of structural simplicity in terms of the presence or not of structure functions. However, since the complete observables scheme may be formulated in terms of the master constraint programme it is arguable that the constraint algebra relevant to it is no more complex. Furthermore, even if we assume that we are dealing with complete observables formulated according to the full Bergmann-Komar algebra (i.e., using the prescription of Dittrich (2006) where the complete observables are constructed with respect to all the constraints) the there is still an issue. The full $\mathcal{BK}$-algebra is undoubtedly more complex than that relevant to shape dynamics because the structure function in the Poisson bracket between two local Hamiltonian constraints has no parallel (refer back to §10.2 and §12.3 to see this explicitly). However, as has been discussed (again in §10.2), it is precisely because of foliation invariance that these structure functions occur. Thus, although tenable, such a line of argument again comes dangerously close to question begging. More broadly, as has been emphasised for the other test cases, there are good historical precedents for the seemingly excess structure of a theory being an essential stepping stone to future development. To jettison the $\mathcal{BK}$-algebra simply on grounds that it is very complex is arguably a rather short sighted move – the fundamental picture of reality we end up with might be foliation invariant *and* scale invariant.

A further potential criterion for privileging one formalism over the other would be a perceived advantage with regard to quantisation. If it could be shown that one formalism provided structures better suited to the application of viable quantisation techniques, then some form of heuristic fruitfulness case could be made for that formalism. Given the technique we introduced above for the relational quantisation of systems with global Hamiltonian constraints in Chapter 9, and the preliminary work of Chapter 15, we might seek to invoke such an argument in favour of shape dynamics. However, as was just mentioned, it is possible that the structure of local Hamiltonian constraints might just as well also prove important in future theoretical development. Moreover, when considered in the context of the master constraint programme the complete observables scheme has also much to offer in terms of potentially heuristically significant structure. In keeping with the preceding discussion, we would argue that for this case of underdetermination we are best served by not giving either formalism an ontologically privileged status and, rather, focusing our attention upon the structural connections that exist between them.

19.3 A problem for ontic structural realism?

We have, then, a formidable yet well-defined philosophical challenge before us. Can we, as in the cases of Lagrangian and Hamiltonian formulations of Newton mechanics and reduced and unreduced formulations of standard gauge theory, find a suitable generalising framework within which to give a structural realist ontology based upon the shape dynamics and compete observables formalisms? This challenge is very much along the same line as that which Pooley (2006) identifies[87] and, as stated above, we concur with him as to the requirement that any adequate structural realist account of structural ontology must provide a 'unifying framework' (although we are not here insisting that this framework need be able to be interpreted 'as corresponding more faithfully to reality than do its various realist representations'.)

In the two cases above we attempted to provide just such a framework in terms of maps between the relevant observables, state spaces and, when necessary, symmetries. Such maps encode the fundamental dynamical and kinematical structure of the theory and thus can be thought of as both suitably 'unifying' and dynamical. This seeming success might drive us to attempt a closely analogous strategy for the case in hand. Thus we might consider the structures relating the observables, state space and symmetries in shape dynamics to the corresponding structures in the complete observables scheme.

With regard to the state spaces and symmetries at least we have a precise mathematical definition of the relevant relationship in terms of the *translation dictionary* that Gomes *et al.* (2011) define between the conventional canonical theory and shape dynamics. This is because in these aspects the complete observables scheme simply is the same as standard canonical theory. However, unlike in our previous two cases it does not seem entirely clear that these maps really do encode anything like the essential structure common be-

[87]The *three* rival formulations of general relativity he challenges the structural realists to account for are: i) Barbour's original 3-space approach; ii) traditional curved spacetime theory; and iii) 'formulations involving spin-2 fields on a flat (or at least fixed) background spacetime'. Our case is analogous to the two way underdetermination of i) vs. ii) rather than Pooley's full three way underdetermination problem.

tween the two formalisms. To provide a means to exchange one symmetry for another is not to provide a basis for why these symmetries can be exchanged in the first place. Furthermore, and relatedly, dynamics in the two cases is represented in an entirely different way and the relevant maps between phase spaces and constraint algebras seemingly encode none of the relevant information.

Of particular importance is the very different way in which the Hamiltonian constraints are both manifested and dealt with within the formalisms. Whereas, the single Hamiltonian constraint of shape dynamics is understood generating dynamical evolution between distinct states of the world, the dynamical aspect of the local Hamiltonian constraints is implemented within the complete observables scheme (including the master constraint version) in a more subtle manner. As detailed above, in the complete observables scheme the Hamiltonian constraints are associated with a flow and the observables are then *smeared* non-locally along this flow. Given such a difference between the formalisms it is difficult to see they could ever be ‘unified’.

Moreover, just as the distinct formal roles played by Hamiltonian constraints within the two approaches would seem irreconcilable so would the algebraic structure of the observables. There is no restriction that observables within the shape dynamics framework must commute with the global Hamiltonian constraint. Thus, although they would be expected to share *some* of the same symmetry properties as complete observables (most importantly being invariant under 3D diffeomorphisms), the shape dynamics observables would form an algebra that cannot be thought of as in any way the same fundamental structure as that of complete observables. In fact, only functions of shape space that are non-dynamic could have correlates amongst the complete observables. The same would be true of any Machian type formulation of canonical general relativity and this is therefore perhaps the most precise concrete realisation of the tension between the two viable strategies for dealing with the problem of time that have informed our discussion. Not only is our case such that the two formalisms are associated, via the appropriate interpretation, with incompatible *ontologies*, but it also seems that, despite the existence of a well-defined translation dictionary, there is a sense in which the *physico-mathematical structures* of the two formulations are themselves incompatible. Thus, as things stand, there is no viable path towards the type of substantive structural realist ontology that we are looking for. In the following final chapter we will consider the implications of this negative result whilst placing it within the wider context of the entire preceding discussion.

Chapter 20

On the interpretation and quantisation of canonical gravity

The principal idea underlying this work has been the dual thesis that the interpretation of classical canonical gravity must be driven by the problem of its quantisation and, conversely, the quantisation of this theory must be driven by its classical interpretation. In this sense, I hope to have defended an attitude towards the philosophy of physics whereby we follow the motto that; *just as the physics must inform the practice of philosophy, so the philosophy of physics should inform the practice of physics*.[88]

The particular application of this strategy that occupied centre stage within Parts I-III was the relationship between interpretative implications of classical geometric reduction and the basis of the Dirac methodology for constraint quantisation. In Part I we considered the sense in which for *standard* gauge theories the interpretation of the classical constraint functions as gauge generating informs their promotion to quantum operators annihilating the wavefunction. We also saw how the structure of quantum gauge theory – including the Faddeev-Popov formulation – can be brought to bear upon debates with regard to *reductionist* interpretations at a classical level.

Part II constitutes the most full implementation of our dual approach towards the practice of physics driven philosophy and philosophy driven physics. We first identified how the mathematical structure of non-relativistic classical reparameterisation invariant theory places restrictions upon the available interpretational stances as to time and possibility space reduction. We then considered the formal and conceptual basis of the two most viable strategies for representing change and observables within the classical theory: the emergent time strategy and the correlation strategy. Next, we utilised our negative argu-

[88] It is not only mathematicians who should remember that, ‘physics is too important to be left to the physicists’ (variously attributed to David Hilbert)

ment with regard to reductionism to motivate a criticism of the application of conventional quantisation techniques to reparameterisation invariant theories. Finally, starting from interpretative basis of Machian temporal relationalism, we were able to find a *relational* technique for quantisation of the theories in question.

In Part III our treatment was extended to the full canonical gravity case. Much of the added subtlety within the relativistic problem of time can be traced to the complex and dual role played by the local Hamiltonian constraints of the theory. In contrast to the non-relativistic case we cannot straightforwardly treat these constraints as purely generating evolution. However, like in the case of non-relativistic models, it can be shown that to simply reduce out the action of the constraint is to trivialise our dynamics. Thus, to move to a denial of time in the interpretation of canonical gravity on the grounds of such reductionism is an incoherent step. There are, however, two alternative stances which do amount to viable interpretational strategies for denying time – the first of which involves removing absolute scale but recovering an emergent notion of time (the Machian denial) and the other which keeps scale but dispenses with change altogether (the Paramenadian denial). These two denials correspond to the emergent time and correlation strategies of the non-relativistic case.

In the penultimate chapter of Part III we argued *from* our interpretative stance with regard to the conceptual failure of reductionism *to* a prescription against the application of conventional constraint quantisation techniques to canonical general relativity. Quantum Hamiltonian constraints cannot be treated like normal constraints because we cannot interpret their classical counterparts as purely gauge generating. The next step of this analysis is to apply the relativistic application of the relational quantisation technique introduced in Part II. Given either the single global Hamiltonian constraint of shape dynamics or the possibility of combining relational quantisation with the master constraint programme, this avenue appears potentially highly fruitful. Work towards the first option was included in Chapter 15.

The early sections of Part IV were principally concerned with introducing several key ideas from the philosophy of science. In particular, the semantic conception of theory structure, the apparent conflict between metaphysical underdetermination and scientific realism, and the ideas surrounding the position of ontic structural realism (OSR). Two points that were much emphasised were that a structuralist ontology must be: i) substantial enough to be a generalising framework which includes dynamical as well as mathematical structure; and ii) such that its essential elements are consistent between two formulations of one theory and between a theory and its successor.

In light of the second of these requirements we may make the proposal that we could test the OSR programme via the investigation of situations where we have two (suitably metaphysically underdetermining) formulations of a classical theory that we can then compare with their quantum analogues. If the structuralist notion of ontology is a coherent one then we should be able to take the frameworks we use to generalise formulations at the classical-classical and quantum-quantum levels and make a further classical-quantum generalisation. These ideas were applied concretely with a degree of success for the cases of Lagrangian and Hamiltonian formulations Newtonian mechanics and the reduced and unreduced formulations of a canonical standard gauge theory.

Finally we come back to the point at which we left off in the last chapter. Upon investigation it is found that there are large – possibly insurmountable – obstacles to the construction of a suitable generalising structuralist framework for canonical general relativity as formulated in terms of shape dynamics and the complete observables scheme. Unlike in our other two case studies there is no suitable set of maps that encodes the structural relationship between the observables, symmetries and dynamics fundamental to the two formalisms (or at least no such set of maps has yet been identified). We are yet to comment on what implications should be attached to this result. Let us review the most obvious options. First, it could be taken that we have found evidence of weakness in the structural realist notion of ontology – in this case it is not fully applicable. We might therefore, assuming we still want to be realists about something, simply fall back to privileging the ontology associated with one formulation or the other – albeit without a totally rational basis for doing so. In practice, this seems closer to what physicists working on this problem are actually doing. A second option would be to reject the idea that the structural ontology we should be looking for in this case must be as substantive as that constructed for the two other cases. There is still a duality between the two formulations and therefore we still have available *some* structural bridges. The problem with such a move is that it exposes the structural realist notion of ontology to twin dangers of triviality and ad-hocness. If all that were needed to constitute a structural ontology were some set of maps then we would, by definition, always be able to satisfy this requirement whenever we had two formulations of a theory. Similarly problematically, if we allow the notion of a substantive structural ontology to be weakened or adjusted whenever we are confronted with a recalcitrant case, then it might seem we are merely adding epicycles onto a degenerating philosophical research programme.

The third option, which arguably has the most promise, is to place this problem of identifying the classical structure of canonical gravity in the context of the outstanding problem of quantising the theory. Within classical theories where we *can* identify the fundamental classical structure we find that this structure is then key to quantisation. Given that we have a situation where, despite over half a century of work, the quantisation of the canonical gravity is *still* seemingly beyond us, would it not best behove us to think of our structuralism problem as part and parcel of the same issue? More specifically, the identification of the *true* degrees of freedom of a classical gauge theory allows us in principle to construct a reduced phase space which correctly parameterises the fundamental dynamics of the theory. It is the symplectic structure of this space which is analogous to the inner product structure of the fundamental Hilbert space of the quantum theory and which is directly linked to the binary operation which defines the algebra of quantum observables. Furthermore, the relationship between the formalism in this space and the unreduced formalism is exactly what we identified above as encoding what is essential to a theory at both classical and quantum levels – states spaces, symmetries, observables, dynamics and the structures which connect them. As we have seen, for canonical gravity this reduced space is *not* the space reached by application of conventional constraint reduction methods and therefore, from the vantage of hindsight, it is perhaps no surprise, that quantisation along conventional lines has failed to deliver a completed theory of quantum gravity.

If, however, we were able to correctly isolated the *true* reduced phase space of canonical gravity (in analogy to what was done for Jacobi's theory in Part II) then we may be better placed to proceed towards quantisation. Furthermore, our expectation would be that the identification of the true reduced space of the canonical theory should allow us to better understand both the complete observables scheme *and* its connection to shape dynamics. One would expect, in fact, that there should be a shape dynamics analogue to this true reduced phase space. More speculatively, we might even propose that it is the connection between these putative reduced spaces that would constitute exactly the substantive structural framework that we are looking for. Thus, in the task of making sense of the interpretation of canonical gravity in terms of structural realism, we might – to recall our epigraph – find that what may have seemed like a tripwire is, in fact, a new path to quantum gravity.

Bibliography

Abraham, R. and Marsden, J. E. (1978). *Foundations of Mechanics*. Benjamin/Cummings.

Albeverio, S. A., Høegh-Krohn, R. J. and Mazzucchi, S. (2008). *Mathematical Theory of Feynman Path Integrals*. Springer-Verlag.

Anderson, E. (2010). The problem of time in quantum gravity.
URL: *arxiv.org/abs/1009.2157*

Anderson, E., Barbour, J., Foster, B. Z., Kelleher, B. and Ó'Murchadha, N. (2005). The physical gravitational degrees of freedom. *Classical and Quantum Gravity* **22**: 1795–1802. doi:10.1088/0264-9381/22/9/020. arXiv:.

Ariew, R. (2011). Pierre Duhem.
URL: *plato.stanford.edu/entries/duhem/*

Arnold, V. I., Kozlov, V. V. and Neishtadt, A. (1988). *Mathematical aspects of celestial mechanics*. Springer-Verlag.

Arnowitt, R., Deser, S. and Misner, C. W. (1960). Canonical variables for general relativity. *Phys. Rev.* **117**: 1595–1602. doi:10.1103/PhysRev.117.1595.
URL: *http://link.aps.org/doi/10.1103/PhysRev.117.1595*

——— (1962). The Dynamics of General Relativity. *In* L. Witten (ed.), *Gravitation: An Introduction to Current Research*, John Wiley & Sons Inc., New York, chapter 7, pp. 227–265.

Babelon, O. and Viallet, C. M. (1979). The geometrical interpretation of the Faddeev-Popov determinant. *Physics Letters B* **85**(2-3): 246 – 248.

Baez, J. C. and Muniain, J. P. (1994). *Gauge fields, knots, and gravity*. World Scientific.

Baierlein, R. F., Sharp, D. H. and Wheeler, J. A. (1962). Three-Dimensional Geometry as a Carrier of Information about Time. *Phys. Rev.* **126**: 1864–1865. doi:10.1103/PhysRev.126.1864.

Bain, J. (2009). Towards structural realism.
URL: *http://philpapers.org/rec/BAITSR*

Barbour, J. (1994). The timelessness of quantum gravity: I. The evidence from the classical theory. *Class. Quant. Grav.* **11**(12): 2853–2873.

——— (1995). General Relativity as a Perfectly Machian Theory. *In* J. B. Barbour and H. Pfister (eds.), *Mach's Principle: From Newton's Bucket to Quantum Gravity*. pp. 214–236.

——— (2000). *The end of time*. Oxford University Press, paperback edition.

——— (2009). Mach's principle, general relativity and gauge theory (unpublished draft).

Barbour, J., Foster, B. and Murchadha, N. O. (2002). Relativity without relativity. *Classical and Quantum Gravity* **19**: 3217.

Barbour, J. and Foster, B. Z. (2008). Constraints and gauge transformations: Dirac's theorem is not always valid. *ArXiv e-prints* arXiv:0808.1223.

Barbour, J. and O'Murchadha, N. (1999). Classical and quantum gravity on conformal superspace arXiv:gr-qc/9911071.

Barbour, J. and Ó'Murchadha, N. (2010). Conformal Superspace: the configuration space of general relativity. *ArXiv e-prints* arXiv:1009.3559.

Barbour, J. B. and Bertotti, B. (1982). Mach's principle and the structure of dynamical theories. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **382**(1783): 295–306.

Belot, G. (1999). Rehabilitating relationalism. *International Studies in the Philosophy of Science* **13**: 35–52.

——— (2000). Geometry and motion. *British Journal for the Philosophy of Science* **51**: 561–95.

——— (2003). Symmetry and gauge freedom. *Studies In History and Philosophy of Modern Physics* **34**: 189–225.

——— (2007). The representation of time and change in mechanics. *In* J. Butterfield and J. Earman (eds.), *Handbook of Philosophy of Physics*, Elsevier, chapter 2.
URL: *http://philsci-archive.pitt.edu/2549/*

Belot, G. and Earman, J. (1999). From metaphysics to physics. *In* J. Butterfield and C. Pagnois (eds.), *Fron Physics to Philosophy*, Cambridge University Press, pp. 166–186.

——— (2001). Pre-socratic quantum gravity. *In* C. Callender and N. Hugget (eds.), *Physics Meets Philosophy at the Planck Scale*, Cambridge University Press.

Bergmann, P. (1949). Non-linear field theories. *Physical Review* **75**(4): 680–685. doi:10.1103/PhysRev.75.680.

Bergmann, P. G. (1961). Observables in General Relativity. *Rev. Mod. Phys.* **33**: 510–514.

Bojowald, M., Hoehn, P. A. and Tsobanjan, A. (2011). An Effective approach to the problem of time. *Class.Quant.Grav.* **28**: 035006. doi:10.1088/0264-9381/28/3/035006. arXiv:1009.5953.

Brown, J. D. and York, J. W. J. (1989). Jacobi's Action And The Recovery Of Time In General Relativity. *Phys. Rev.* **D40**: 3312–3318. doi:10.1103/PhysRevD.40.3312.

Butterfield, J. (1988). Albert Einstein meets David Lewis. *Proceedings of the Biennial Meeting of the Philosophy of Science Association* **2**: 66–81.

——— (2002). Critical notice. *The British Journal for the Philosophy of Science* **53**(2): 289–330.

——— (2007). On symplectic reduction in classical mechanics. *In* J. Butterfield and J. Earman (eds.), *Philosophy of Physics*, Elsevier, pp. 1–131.

Conrady, F. and Freidel, L. (2009). Quantum geometry from phase space reduction. *Journal of Mathematical Physics* **50**(12): 123510–123510–29.

Corichi, A. (2008). On the geometry of quantum constrained systems. *arXiv* **gr-qc**.
URL: *http://arxiv.org/abs/0801.1119v2*

Daughton, A., Louko, J. and Sorkin, R. D. (1993). Initial conditions and unitarity in unimodular quantum cosmology arXiv:gr-qc/9305016.

DeWitt, B. (1967). Quantum theory of gravity. I. The canonical theory. *Physical Review* **160**: 1113–1148.

Dirac, P. A. (1950). Generalized hamiltonian dynamics. *Can. J. Math.* **2**: 129 – 148.

——— (1959). Fixation of Coordinates in the Hamiltonian Theory of Gravitation. *Physical Review* **114**: 924–930. doi:10.1103/PhysRev.114.924.

Dirac, P. A. M. (1933). The Lagrangian in Quantum Mechanics. *In* J. Schwinger (ed.), *Selected Papers On Quantum Electrodynamics*, Dover Publications, pp. 312–320.

——— (1958a). Generalized Hamiltonian Dynamics. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **246**: 333–3343.

——— (1958b). The Theory of Gravitation in Hamiltonian Form. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **246**: 333–343.

——— (1964). *Lectures on quantum mechanics*. Dover Publications.

Dittrich, B. (2006). Partial and complete observables for canonical general relativity. *Classical and Quantum Gravity* **23**: 6155.

——— (2007). Partial and complete observables for Hamiltonian constrained systems. *General Relativity and Gravitation* **39**: 1891.

Dittrich, B. and Thiemann, T. (2006). Testing the master constraint programme for loop quantum gravity: I. General framework. *Classical and Quantum Gravity* **23**: 1025–1065.

Dowker, F., Johnston, S. and Sorkin, R. D. (2010). Hilbert spaces from path integrals. *Journal of Physics A: Mathematical and Theoretical* **43**(27): 275302.
URL: *http://stacks.iop.org/1751-8121/43/i=27/a=275302*

Duistermaat, H., Guillemin, V., Meinrenken, E. and Wu, S. (1995). Symplectic reduction and Riemann-Roch for circle actions. *Math. Res. Letters* **2**(259-66).

Earman, J. (2003). Tracking down gauge: an ode to the constrained Hamiltonian formalism. *In* K. Brading and E. Castellani (eds.), *Symmetries in physics*, Cambridge University Press, pp. 150–162.

Echeverria-Enriquez, A. and Munoz-Lecanda, M. (1999). Mathematical foundations of geometric quantization.
URL: *http://arxiv.org/abs/math-ph/9904008*

Einstein, A. (1916). Die Grundlage der Allgemeinen Relativitätstheorie. *Annalen der Physik* **49**(769-822).

Faddeev, L. D. (1969). The Feynman integral for singular Lagrangians. *Theoretical and Mathematical Physics* **1**(1): 1–13.

Faddeev, L. D. and Popov, V. N. (1967). Feynman diagrams for the Yang-Mills field. *Physics Letters B* **25**: 29–30.

Feynman, R. P. (1948). Space-Time Approach to Non-Relativistic Quantum Mechanics. *Reviews of Modern Physics* **20**: 367–387.

French, S. (2011). Metaphysical underdetermination: why worry? *Synthese* **180**(2): 205–221.

French, S. and Krause, D. (2006). *Identity in physics: A historical, philosophical and formal account.* Oxford University Press.

Gerlach, U. H. (1969). Derivation of the ten einstein field equations from the semiclassical approximation to quantum geometrodynamics. *Phys. Rev.* **177**: 1929–1941. doi:10.1103/PhysRev.177.1929.

Geroch, R. (1970). Domain of Dependence. *Journal of Mathematical Physics* **11**: 437–449.

Giddings, S. B., Marolf, D. and Hartle, J. B. (2006). Observables in effective gravity. *Phys.Rev.* **D74**: 064018. doi:10.1103/PhysRevD.74.064018. arXiv:hep-th/0512200.

Giulini, D. (2009). The Superspace of geometrodynamics. *General Relativity and Gravitation* **41**: 785–815.

Giulini, D. and Marolf, D. (1999a). A uniqueness theorem for constraint quantization. *Classical and Quantum Gravity* **16**: 2489–2505.

——— (1999b). On the generality of refined algebraic quantization. *Classical and Quantum Gravity* **16**: 2479–2488.

Gomes, H., Gryb, S. and Koslowski, T. (2011). Einstein gravity as a 3D conformally invariant theory. *Classical and Quantum Gravity* **28**(4): 045005–+. doi:10.1088/0264-9381/28/4/045005. arXiv:1010.2481.

Gomes, H., Gryb, S., Koslowski, T. and Mercati, F. (2011). The gravity/CFT correspondence arXiv:1105.0938.

Goodman, N. (1955). *Fact, Fiction, and Forecast, Cambridge,*. Harvard University Press.

Gotay, M., Nester, J. and Hinds, G. (1978). Presymplectic manifolds and the Dirac-Bergmann theory of constraints. *Journal of Mathematical Physics* **19**: 2388–2399.

Gotay, M. J. (1980). Functorial geometric quantization and van hove's theorem. *International Journal of Theoretical Physics* **19**(2): 139–161. doi:10.1007/BF00669766.

Gotay, M. J. (1986). Constraints, reduction, and quantization. *Journal of Mathematical Physics* **27**: 2051–2066.

Gryb, S. (2011). *Shape dynamics and Mach's principles: Gravity from conformal geometrodynamics*. Ph.D. thesis, Univeristy of Waterloo.

Gryb, S. and Thébault, K. (2012). The role of time in relational quantum theories. *Foundations of Physics* **DOI: 10.1007/s10701-012-9665-5**.
URL: *http://arxiv.org/abs/1110.2429*

Gryb, S. B. (2010). A Definition of Background Independence. *Class. Quant. Gravity.* **27**: 215018. doi:10.1088/0264-9381/27/21/215018. arXiv:1003.1973.

Guillemin, V. and Sternberg, S. (1982). Geometric quantization and multiplicities of group representations. *Inventiones Mathematicae* **67**(3): 515–538.

Hardin, C. L. and Rosenberg, A. (1982). In defence of convergent realism. *Philosophy of Science* **49**: 604–15.

Hempel, C. and Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science* **15**: 135–175.

Henneaux, M. and Teitelboim, C. (1989). The Cosmological Constant And General Covariance. *Phys. Lett.* **B222**: 195–199. doi:10.1016/0370-2693(89)91251-3.

——— (1992). *Quantization of gauge systems*. Princeton University Press.

Hilgevoord, J. (2005). Time in quantum mechanics: a story of confusion. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* **36**(1): 29 – 60.

Hojman, S. A., Kuchař, K. and Teitelboim, C. (1976). Geometrodynamics regained. *Annals of Physics* **96**: 88–135. doi:10.1016/0003-4916(76)90112-3.

Isham, C. (1992). Canonical quantum gravity and the problem of time. *Arxiv preprint gr-qc* .
URL: *http://arxiv.org/abs/grqc/9210011*

Isham, C. and Kuchař, K. (1985a). Representations of spacetime diffeomorphisms. I. Canonical parametrized field theories. *Annals of Physics* **164**(2): 288 – 315. doi:DOI: 10.1016/0003-4916(85)90018-1.
URL: *http://www.sciencedirect.com/science/article/pii/0003491685900181*

——— (1985b). Representations of spacetime diffeomorphisms. II. Canonical geometrodynamics. *Annals of Physics* **164**(2): 316 – 333. doi:DOI: 10.1016/0003-4916(85)90019-3.
URL: *http://www.sciencedirect.com/science/article/pii/0003491685900193*

Johns, O. (2005). *Analytical mechanics for relativity and quantum mechanics*. Oxford University Press.

Jones, R. (1991). Realism about what? *Philosophy of Science* **58**(2): 185–202.

Katzav, J. (2004). Dispositions and the principle of least action. *Analysis* **64**(283): 206–214. doi:10.1111/j.0003-2638.2004.00487.x.
URL: *http://dx.doi.org/10.1111/j.0003-2638.2004.00487.x*

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science* **48**: 507–531.

Kuchař, K. (1991a). Does an unspecified cosmological constant solve the problem of time in quantum gravity? *Phys. Rev.* **D43**: 3332–3344. doi:10.1103/PhysRevD.43.3332.

——— (1991b). The problem of time in canonical quantization of relativistic systems. *In* A. Ashtekar and J. Stachel (eds.), *Conceptual Problems of Quantum Gravity*, Boston University Press, p. 141.

——— (1992). Time and interpretations of quantum gravity. *In 4th Canadian Conference on General Relativity* . . . . World Scientific Singapore.

Ladyman, J. (1998). What is structural realism? *Studies In History and Philosophy of Science Part A* **29**(3): 409 – 424.

——— (2007). Scientific structuralism: On the identity and diversity of objects in a structure. *Aristotelian Society Supplementary Volume* **81**: 23.

Ladyman, J. and Ross, D. (2007). *Every thing must go*. Oxford University Press.

Lanczos, C. (1970). *The Variational Principles of Mechanics*. Dover Publications, New York.

Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science* **48**: 19–49.

Lewis, D. (1983). Individuation by acquaintance and by stipulation. *The Philosophical Review* **92**: 3–32.

Lusanna, L. (2011). Canonical gravity and relativistic metrology: from clock synchronization to dark matter as a relativistic inertial effect arXiv:1108.3224v1.
URL: *http://arxiv.org/abs/1108.3224v1*

Mach, E. (1960). *The Science of Mechanics: A Critical and Historical Account of its Development*. Trans. McCormack, T. J. Open Court Publishing Company, LaSalle, Illinois.

Markosian, N. (2011). Time.
URL: *http://plato.stanford.edu/entries/time/*

Marolf, D. (1995). Almost ideal clocks in quantum cosmology: A Brief derivation of time. *Class.Quant.Grav.* **12**: 2469–2486. doi:10.1088/0264-9381/12/10/007. arXiv:gr-qc/9412016.

——— (2009). Solving the Problem of Time in Mini-superspace: Measurement of Dirac Observables. *Phys.Rev.* **D79**: 084016. doi:10.1103/PhysRevD.79.084016. arXiv:0902.1551.

Marsden, J. and Ratiu, S. (1994). *Introduction to mechanics and symmetry*. Springer-Verlag.

Marsden, J. E. and Weinstein, A. (1982). The Hamiltonian structure of the Maxwell-Vlasov equations. *Physica D: Nonlinear Phenomena* **4**(3): 394 – 406.

Maudlin, T. (2002). Thoroughly Muddled McTaggart: Or, How to Abuse Gauge Freedom to Create Metaphysical Monostrosities. *Philosopher's Imprint* **2**(4): 1–19.

Misner, C. W. (1957). Feynman quantization of general relativity. *Rev. Mod. Phys.* **29**(3): 497–509. doi:10.1103/RevModPhys.29.497.

Mittelstaedt, P. (1976). *Der Zeitbegriff in der Physik*. B.I.-Wissenschaftsverlag, Mannheim, Germany.

Nabokov, V. (1955). *Lolita*. Olympia Press.

Nerlich, G. (1979). What can geometry explain. *British Journal for the Philosophy of Science* **30**: 69–83.

Newton, I. (1962). *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and his System of the World*. Trans. Motte, A., revised Cajori, F. University of California Press, Berkeley.

North, J. (2009). The "structure" of physics: A case study? *The Journal of Philosophy* **106**: 57–88.

Norton, J. (2011). The hole argument. *Stanford Encyclopedia of Philosophy* .
URL: *http://plato.stanford.edu/entries/spacetime-holearg/*

Ordóñez, C. R. and Pons, J. M. (1992). Equivalence of reduced, Polyakov, Faddeev-Popov, and Faddeev path-integral quantization of gauge theories. *Phys. Rev. D* **45**(10): 3706–3712. doi:10.1103/PhysRevD.45.3706.

Parrinello, C. and Jona-Lasinio, G. (1990). A modified Faddeev-Popov formula and the Gribov ambiguity. *Physics Letters B* **251**(1): 175 – 180. doi:DOI: 10.1016/0370-2693(90)90249-6.
URL: *http://www.sciencedirect.com/science/article/pii/037026939090249б*

Peskin, M. and Schroeder, D. V. (1995). *An introduction to quantum field theory*. Westview Press.

Pons, J., Salisbury, D. and Shepley, L. (1997). Gauge transformations in the Lagrangian and Hamiltonian formalisms of generally covariant theories. *Phys. Rev. D* **55**(2): 658–668.

Pons, J., Salisbury, D. and Shepley, L. C. (1999). Reduced phase space: quotienting procedure for gauge theories. *Journal of Physics A: Mathematical and General* **32**: 419.

Pons, J., Salisbury, D. and Sundermeyer, K. A. (2010). Observables in classical canonical gravity: folklore demystified. *Journal of Physics A: Mathematical and General* **222**(12018).

Pooley, O. (2001). Relationism Rehabilitated? II: Relativity.
URL: *http://philsci-archive.pitt.edu/221/*

——— (2006). Points, particles and structural realism. *In* S. French, D. Rickles and J. Saatsi (eds.), *Structural Foundations of Quantum Gravity*, Oxford University Press, pp. 83–120.

Pooley, O. and Brown, H. R. (2001). Relationalism Rehabilitated? I: Classical Mechanics. *British Journal for the Philosophy of Science* **53**: 183.

Popov, V. N. (2010). Faddeev-Popov ghosts.
URL: *http://www.scholarpedia.org/article/Faddeev-Popovghosts*

Popov, V. N. and Faddeev, L. D. (2005). Perturbation theory for gauge-invariant fields. *In* G. 't Hooft (ed.), *50 Years of Yang-Mills Theory*, World Scientific.

Price, H. (2009). The flow of time.
URL: *http://philsci-archive.pitt.edu/4829/*

Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth*. London: Routledge.

Redhead, M. (2003). The interpretation of gauge symmetry. *In* K. Brading and E. Castellani (eds.), *Symmetries in physics*, Cambridge University Press, pp. 124–139.

Redhead, M. L. G. (1999). Quantum field theory and the philosopher. *In* T. Cao (ed.), *Conceptual Foundations of Quantum Field Theory.*, Cambridge University Press.

Rickles, D. (2008). *Symmetry, Structure, and Spacetime*, volume 3 of *Philossphy and Foundations of Physics*. Elsevier.

Rovelli, C. (1990). Quantum mechanics without time: A model. *Phys. Rev. D* **42**: 2638–2646.

Rovelli, C. (1991). Time in quantum gravity: An hypothesis. *Phys. Rev. D* **43**: 442.

——— (2002a). Gps observables in general relativity". *Phys. Rev. D; Physical Review D* **65**(4). doi:10.1103/PhysRevD.65.044017.

——— (2002b). Partial observables. *Phys. Rev. D* **65**: 124013.

——— (2004). *Quantum Gravity*. Cambridge University Press, Cambridge.

——— (2007). Comment on "Are the spectra of geometrical operators in Loop Quantum Gravity really discrete?" by B. Dittrich and T. Thiemann.
URL: *http://arxiv.org/abs/0708.2481*

Salisbury, D. C. (2007). Rosenfeld, Bergmann, Dirac and the Invention of Constrained Hamiltonian Dynamics. arXiv:physics/0701299v1.
URL: *http://arxiv.org/abs/physics/0701299v1*

——— (2010). Léon Rosenfeld's pioneering steps toward a quantum theory of gravity. *Journal of Physics: Conference Series* **222**(1): 012052.
URL: *http://stacks.iop.org/1742-6596/222/i=1/a=012052*

——— (2012). Peter Bergmann and the Invention of Constrained Hamiltonian Dynamics.

Saunders, S. (1993). To what physics corresponds. *In* H. Kaminga and S. French (eds.), *Correspondence, invariance, and heuristics; essays in honour of heinz post*, Dordrecht: Kluwer, pp. 295–326.

——— (2003). Physics and Leibniz's Principles. *In* K. Brading and E. Castellani (eds.), *Symmetries in physics*, Cambridge University Press, pp. 289–307.

Smart, B. (2012). *A Critique of Humean and Anti-Humean Metaphysics of Cause and Law*. Ph.D. thesis, Univeristy of Nottingham.

Smeenk, C. and Wuthrich, C. (2009). Time travel and time machines. Forthcoming in C. Callender (ed.), Oxford Handbook of Time, Oxford University Press.
URL: *http://philsci-archive.pitt.edu/4915/*

Smolin, L. (2009). The quantization of unimodular gravity and the cosmological constant problem arXiv:0904.4841.

Sorkin, R. D. (1994). Quantum mechanics as quantum measure theory. arXiv:gr-qc/9401003v2.
URL: *http://arxiv.org/abs/gr-qc/9401003v2*

——— (1997). Forks in the Road, on the Way to Quantum Gravity. *Int. J. Theor. Phys.* **36**: 2759–2781. doi:10.1007/BF02435709. arXiv:gr-qc/9706002.

Souriau, J. (1997). Structure of dynamical systems: a symplectic view of physics. *In* R. Cushman and G. Tuynman (eds.), *Cushman-de Vries, C. H (trans.)*, Birkhauser.

Sundermeyer, K. (1982). *Constrained dynamics: with applications to Yang-Mills theory, general relativity, classical spin, dual string model*. Lecture notes in physics. Springer-Verlag.
URL: *http://books.google.com/books?id=Wvx8QgAACAAJ*

Swinburne, R. (1997). *Simplicity as Evidence for Truth*. Marquette University Press.

Teitelboim, C. (1973). How commutators of constraints reflect the spacetime structure. *Annals of Physics* **79**(2): 542 – 557. doi:DOI: 10.1016/0003-4916(73)90096-1.
URL: *http://www.sciencedirect.com/science/article/pii/0003491673900961*

Teller, P. (1995). *An interpretive introduction to quantum field theory*. Princeton University Press.

Thébault, K. P. Y. (2011). Quantisation, representation and reduction; how should we interpret the quantum hamiltonian constraints of canonical gravity? *Symmetry* **3**: 134.

——— (2012). Symplectic reduction and the problem of time in nonrelativistic mechanics. *British Journal for the Philosophy of Science (doi: 10.1093/bjps/axr045, print verison forthcoming)* .

Thiemann, T. (2006). The phoenix project: master constraint programme for loop quantum gravity. *Classical and Quantum Gravity* **23**: 2211.

——— (2007). *Modern canonical quantum general relativity*. Cambridge University Press.

Thompson, P. (1989). *The Structure of Biological Theories*. State University of New York Press.

Unruh, W. G. (1989). A Unimodular Theory Of Canonical Quantum Gravity. *Phys. Rev.* **D40**: 1048. doi:10.1103/PhysRevD.40.1048.

Unruh, W. G. and Wald, R. M. (1989). Time And The Interpretation Of Canonical Quantum Gravity. *Phys. Rev.* **D40**: 2598. doi:10.1103/PhysRevD.40.2598.

Unz, R. K. (1986). Path integration and the functional measure. *Il Nuovo Cimento A (1971-1996)* **92**(4): 397–426.

van Fraassen, B. C. (1970). On the Extension of Beth's Semantics of Physical Theories. *Philosophy of Science* **37**(3).

——— (1980). *The scientific image*. Oxford University Press.

van Fraassen, B. C. (2008). *Scientific representation: paradoxes of perspective*. Oxford University Press.

Walters, P. (1981). *An Introduction to Ergodic Theory*. Springer-Verlag: New York.

Westenholz, C. (1978). *Differential forms in mathematical physics*. North-Holland.

Westman, H. and Sonego, S. (2008). Events and observables in generally invariant spacetime theories. *Foundations of Physics* **38**(10): 908–915. doi:10.1007/s10701-008-9235-z.

Weyl, H. (1918). Gravitation und elektrizitat. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin* pp. 465–480.

——— (1922). *Space-time-matter*. Dover, New York, NY. Translation by Brose, Henry L from the original German text into English.

——— (1939). *The classical groups: their invariants and representations*. Princeton University Press.

Wheeler, J. A. (1968). Superspace and the Nature of Quantum Geometrodynamics. *In* J. A. Wheeler and C. M. De Witt (eds.), *Lectures in Mathematics and Physics*, Benjamin.

Woodhouse, N. (1997). *Geometric Quantization*. Oxford Mathematical Monographs. Clarendon Press.
URL: *http://books.google.com/books?id=PuhpI98ZkDgC*

Woodward, J. (2009). Scientific explanation.
URL: *http://plato.stanford.edu/entries/scientific-explanation/*

Worrall, J. (1989). Structural realism: The best of both worlds? *Dialectica* **43**(1-2): 99–124.

Wüthrich, C. (2010). No presentism in quantum gravity. *In Space, Time, and Spacetime*, volume 167 of *Fundamental Theories of Physics*, pp. 257–278.

York, J. J. W. (1972). Role of conformal three geometry in the dynamics of gravitation. *Phys. Rev. Lett.* **28**: 1082–1085. doi:10.1103/PhysRevLett.28.1082.

York, J. W., Jr. (1973). Conformally invariant orthogonal decomposition of symmetric tensors on Riemannian manifolds and the initial-value problem of general relativity. *Journal of Mathematical Physics* **14**: 456–464. doi:10.1063/1.1666338.