SHAPE DYNAMICS

RELATIVITY AND RELATIONALISM

FLAVIO MERCATI

OXFORD

SHAPE DYNAMICS

Shape Dynamics

Relativity and Relationalism

Flavio Mercati

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2018
Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2017959949

ISBN 978–0–19–878947–5 (hbk.)
ISBN 978–0–19–878948–2 (pbk.)

DOI 10.1093/oso/9780198789475.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Cover picture concept by J. Barbour

'From henceforth the planets follow their paths through the ether like the birds in the air. We must therefore philosophize about these things differently.'
Johannes Kepler (1571–1630)

Contents

1 Introduction 1

1.1 Outline of Chapters 1
1.2 Notation 4
1.3 Shape Dynamics in a nutshell 4

Part I Historical Motivation

2 Newton's Bucket 17

2.1 The defects of the law of inertia 17
2.2 Leibniz's relationalism 18
2.3 The *Scholium* problem 19
Further reading 20

3 Origins of the Mach–Poincaré Principle 21

3.1 Tait's partial solution of the Scholium problem 21
3.2 Mach's critique of Newton 23
3.3 Hoffman's experiment 25
3.4 Poincaré's principle 26
Further reading 27

Part II Relational Particle Dynamics

4 Barbour–Bertotti Best Matching 31

4.1 Best matching: intuitive approach 31
Further reading 38

5 Best Matching: Technical Details 39

5.1 Mathematical Formulation 39
5.2 Temporal relationalism: Jacobi's principle 44
5.3 Best matching 'in action' 48
Further reading 55

6 Hamiltonian Formulation 56

6.1 The Hamiltonian constraint 56
6.2 A crash course in Dirac's constraint analysis 57
6.3 Application to our systems 61
6.4 A matter of units 63
Further reading 66

Part III Relational Field Theory

III.1 The Ontology of Fields 67
III.2 The origins of Geometrodynamics 69
III.3 An exercise in counterfactual history 70
III.4 Nuggets of functional analysis 71

7 Relativity Without Relativity 73

7.1 Rigidity of the choice of the Lagrangian 76
7.2 Inclusion of a scalar field: Special Relativity 77
7.3 Inclusion of a one-form field: Gauge Theory 80
7.4 Best-matching gauge transformations 82
7.5 Yang–Mills Theory 83
7.6 Further generalizations 84
Further reading 85
7.7 The problem of many-fingered time 85

8 York's Solution to the Initial-Value Problem 88

8.1 Lichnerowicz's partial solution 89
8.2 York's general solution 92
Further reading 94

9 A derivation of Shape Dynamics 95

9.1 Analysis of Case III 98
Further reading 102

10 Cotton-Squared Theory 103

Part IV Shape Dynamics

11 Historical Interlude 109

12 Shape Dynamics and the Linking Theory 112

12.1 Linking Theory's equation of motion 114
12.2 The degrees of freedom of Shape Dynamics 117
12.3 The solution to the problem of time in SD 117
12.4 Coupling to matter and uniqueness 122
12.5 Experienced spacetime 124
12.6 'Symmetry doubling': BRST formulation of SD 126
12.7 'Conformal geometrodynamics regained' 128

13 Solutions of Shape Dynamics 133

13.1 Homogeneous solutions: Bianchi IX 134
13.2 Continuation through the singularity 151
13.3 Spherically symmetric solutions 169
13.4 Asymptotically flat Shape Dynamics 209

Part V Appendices

Appendix A Arnowitt–Deser–Misner Gravity 223

A.1 The Arnowitt–Deser–Misner (ADM) formalism 223
Further reading 225
A.2 The Wheeler–DeWitt equation 226
A.3 The Baierlein–Sharp–Wheeler action 226
A.4 Asymptotically flat ADM 227

Appendix B Other Appendices 230

B.1 The case for closed spacelike hypersurfaces 230
B.2 Free-end-point variation 231
B.3 Lie derivative 232
B.4 TT decomposition of tensors 233
B.5 Point sources in ADM gravity 238
B.6 The poles of a spherically-symmetric universe 239

References 243

Index 249

List of Figures

1	What is a Conformal Transformation	5
2	The concept of a gauge constraint	6
3	The concept of a gauge-fixing	7
4	The iconic diagram of Shape Dynamics	9
5	Scheme of the constraints of SD and GR	10
6	Newton's bucket experiment	19
7	Hoffman's experiment	26
8	Best matching illustrated with pictures of stars	36
9	Horizontal stacking	36
10	3-body configuration space as a fibre bundle	41
11	*G*-invariance condition for a connection	42
12	Stage I of best matching	44
13	Stage II of best matching	45
14	Vertical stacking	46
15	Faraday's lines of force	68
16	The issue of Many-Fingered Time	86
17	Graphical identification of the root of the LY polynomial	93
18	2d quantum harmonic oscillator	119
19	The Dirac algebroid	129
20	Shape potential of Bianchi IX	138
21	On-shell surface in Kasner space	139
22	Intersection between Kasner sphere and plane	140
23	Division of Bianchi IX shape space	142
24	Taub transition	144
25	Exponents in the Bianchi IX shape potential	147
26	Bianchi IX shape potential vs. shape kinetic energy	148
27	Bianchi IX shape potential vs. shape kinetic energy, $\Xi < 1$	149
28	Spherical Bianchi IX shape space	151
29	Bianchi I solutions on Bianchi IX shape space	153

30 Quiescence bound in dimensionless variables—1 166
31 Quiescence bound in dimensionless variables—2 167
32 Domain of areal radius 174
33 The surface $\mathscr{P}(z) = 0$ for $m > 0, \lambda < 0$ 179
34 The surface $\mathscr{P}(z) = 0$ for $m > 0, \lambda \geq 0$ 180
35 The surface $\mathscr{P}(z) = 0$ for $m < 0, \lambda > 0$ 181
36 The surface $\mathscr{P}(z) = 0$ for $m = 0, \lambda > 0$ 182
37 Single shell universe diagram 189
38 Solutions of the constraints for the single-shell universe 190
39 $\sigma(R)$ and V vs. $\langle p \rangle$ in single-shell universe 194
40 Twin-shell universe diagram 197
41 On-shell surfaces for $\lambda < 0, m_{\mathrm{B}} > 0$ 200
42 On-shell surfaces for $\lambda > 0, m_{\mathrm{B}} > 0$ 201
43 On-shell surfaces for $\lambda > 0, m_{\mathrm{B}} < 0$ 202
44 On-shell surfaces for $\lambda > 0, m_{\mathrm{B}} = 0$ 203
45 On-shell surface for $\lambda < 0, m_{\mathrm{B}} < 0$ 204
46 On-shell surface for $\lambda < 0, m_{\mathrm{B}} = 0$ 205
47 Determining the codomain of σ in the twin-shell universe 206
48 Determining the codomain of σ in the twin-shell universe 208
49 The neighbours of the solar system 210
50 AF single-shell on-shell surfaces 217
51 Regions of Kruskal's extension occupied by the 'wormhole' solution 218
52 Lapse and Shift 224

Acknowledgements

This book would not have existed without the initial encouragement and continued support of Julian Barbour, to whom I am immensely grateful. I send a big thank to Henrique Gomes for his extensive help at various stages of composition of this book. I am indebted also to my mentor Lee Smolin and the fantastic environment of Perimeter Institute for their support to a young researcher that took the risk of writing a book during a postdoctoral appointment. Thanks also to Tim Koslowski, David Sloan, David Wiltshire and Matteo Lostaglio for their valuable comments and discussions. Special thanks to my former students Andrea Napoletano, Monica Rincon Ramirez and Mykola Murskyj for being my guinea pigs and proofreading early drafts of what became the first monograph on Shape Dynamics.

1 Introduction

Shape Dynamics (SD) is a new theory of gravity that is based on fewer and more fundamental first principles than General Relativity (GR). The most important feature of SD is the replacement of GR's relativity of simultaneity with a more tractable gauge symmetry, namely, invariance under spatial conformal transformations. This book contains both a quick introduction for readers curious about SD and a detailed walk through of the historical and conceptual motivations for the theory, its logical development from first principles, and an in-depth description of its present status. The book is sufficiently self-contained for graduate students and advanced undergraduate students with some basic background in GR and Lagrangian/Hamiltonian mechanics. It is intended both as a reference text for students approaching the subject and as a reviewing tool for researchers interested in the theory.

1.1 Outline of Chapters

The main part of the introduction is Section 1.3 of Chapter 1 (Shape Dynamics in a nutshell), where I attempt to offer a no-nonsense quick entry to the basic ideas of SD. This serves dual purposes: on one hand, students interested in SD will have a brief overview of what the theory is about and what we hope to achieve with it; on the other, researchers curious about SD will find a short description of the theory in Sec. 1.3 that is hopefully enough for them to decide whether these ideas are worth examining in depth. The minimum of notions needed to understand the core ideas of SD are outlined with the aim of making the section as self-contained as possible. All the concepts are explained in detail in the rest of this book, while taking an 'historico-pedagogical' perspective and introducing them at the appropriate points in the story. Sec. 1.3 also includes a quick outline of basic concepts needed to understand SD, which are not part of normal undergraduate curricula (like constrained Hamiltonian systems and gauge theories). However the section is limited to just a few pages to be read quickly by experts, and the mentioned outline is by no means sufficient to understand those concepts properly. Its purpose is to give undergraduate readers a taste of the background knowledge that is necessary to understand SD and get the overall drift. Everything is exhaustively explained in the body of the text.

Shape Dynamics: Relativity and Relationalism. Flavio Mercati.
Published 2018 by Oxford University Press. 10.1093/oso/9780198789475.001.0001

Part I shows where SD comes from: we consider it as the most advanced stage of the *relational programme,* which seeks to eliminate all absolute structures from physics—absolute structures meaning anything that determines physical phenomena but is not determined by them. The chief example is Newton's absolute space and time (or, in modern terms, inertial frames of reference). The battlefield of Newton's absolutes has seen giants of science fighting the absolute-vs-relative debate: Galileo, Descartes, Newton himself, Leibniz, Mach, Poincaré and Einstein. Another example is scale, or size: SD tries to eliminate precisely this absolute structure from physics. One could imagine pushing this programme further into the future: what determines the topology of space? Are the values of the physical constants a result of immutable laws or of a dynamical evolution?

In Chapter 2 I will explain in detail the fundamental problem of Newtonian dynamics: everything is based on the law of inertia, which in turn relies on the concepts of rest and uniform motion, but these concepts are not defined by Newton. Chapter 3 makes it clear what the problem with Newton's construction is. Stating in a mathematically precise way the defect of Newton's theory was an incredibly difficult problem; Henri Poincaré solved it after more than two centuries. However, even Poincaré's formulation (which we call the 'Mach–Poincaré Principle') was not recognized for its true worth until the work of Barbour and Bertotti in the 1960s.

Part II deals with relational dynamics in the simpler framework of systems of point particles. Relational dynamics is a reformulation of dynamics that satisfies the Mach–Poincaré Principle, as formulated by Barbour and Bertotti. It uses specific techniques that were invented for purpose, in particular that of 'best matching'. These techniques turned out to be equivalent to the modern formulation, due to Dirac, of gauge theories as constrained Hamiltonian systems. In Chapter 4 best matching is introduced at an intuitive level, while Chapter 5 details it using the language of Principal Fibre Bundles, which are introduced to the reader. Chapter 6 describes the Hamiltonian formulation of best matching and links it to modern gauge theory. The techniques developed by Dirac for Hamiltonian constrained systems are needed in this section and are therefore briefly explained.

Part III deals with the more advanced framework of field theory. Chapter 7 details (in modern language) a series of results due to Barbour, O'Murchadha, Foster, Anderson and Kelleher. These are striking results: they show that the principles of relational field theory alone are sufficient to derive GR, the general and special relativity principles, the universality of the light cone, Maxwell's electromagnetism, the gauge principle and Yang–Mills theory. Chapter 8 contains more background material: it presents York's method for the solution of the initial-value problem in GR. This provides an important input for the formulation of SD. Chapter 9 deals with work I have done together with E. Anderson, and finally makes the connection from relational field theory to SD. The latter is shown to arise from the principles of relational field theory and the Mach–Poincaré Principle. The final chapter of this part, Chapter 10, describes the attempt to build a theory that incorporates the principles of relational field theory and assumes local scale invariance (also called *conformal*, or *Weyl* invariance) from the beginning (while in my derivation of SD local scale invariance emerges as a consistency requirement in the analysis of the constraints of the theory). Such a theory would implement both a local notion of duration (and therefore invariance under local time reparametrization) and conformal invariance. Interestingly, this theory proves to be inconsistent, leaving us with SD as the only viable candidate for a theory of evolving conformal geometry. As SD only has a global reparametrization invariance and implies a preferred notion of simultaneity, we have to conclude that refoliation invariance and conformal invariance are dual and alternative to each other: they cannot be kept simultaneously.

In Part IV SD is finally formulated in its current form. I begin with a brief account of the way the ideas at the basis of SD were developed in Chapter 11, then in Chapter 12 I proceed to derive the equations of SD from the point I left the theory in Chapter 9. In Sec. 12.1 I shall discuss the physical degrees of freedom of SD, which are the conformally invariant properties of a three-dimensional (3D) manifold, and their conjugate momenta. In Sec. 12.2 I explain how SD represents a simple solution to the problem of time in quantum gravity, and how one reconstructs the familiar four-dimensional spacetime description of GR from a solution of SD. Sec. 12.3 deals with the coupling of SD to matter, which was analized by Gomes and Koslowski, who applied to SD previous results on the conformal method by Isenberg, Nester, Ó Murchadha and York. In Sec. 12.4 I summarize Koslowski's work on the emergence of the spacetime description. This work shows how the four-dimensional, CMC-foliated line element that one deduces from a solution of SD is the spacetime that matter degrees of freedom experience in the limit in which backreaction can be ignored. In Sec. 12.5 I briefly describe the results by Koslowski and Gomes on the BRST formulation of SD, and finally in Section 12.6 I summarize Gomes' work on a construction principle for SD along the line of rigidity theorems like those of Hojman-Kuchař-Teitelboim.

Chapter 13 deals with the particular solutions of SD that have been studied so far. In Sec. 13.1 I study homogeneous solutions in detail, with spherical topology (the so-called 'Bianchi IX' universes), and show what is perhaps the most striking consequence of SD: its solutions can be continued uniquely through the Big Bang singularity. In Sec. 13.3 I analyze spherically-symmetric solutions, which are the basis for discussing gravitational collapse and black holes, and present another striking result: the Arnowitt–Deser–Misner (ADM)-in-constant-mean-extrinsic curvature (CMC)-foliation description of a closed universe with collapsing matter fails at some point during the collapse (presumably when the system generates an event horizon), while the SD description seems well-defined at that point and afterwards. In Sec. 13.4 I discuss the sense in which one can talk about asymptotic flatness in SD (which is fundamentally a theory of compact universes), and I critically evaluate past results obtained in the asymptotically flat case.

The final part of the book contains the appendices, which are divided into a first, major Appendix A, with a brief account of the Hamiltonian formulation of GR due to Arnowitt, Deser and Misner. This is the main tool of Canonical General Relativity and is the theory we have to compare classical SD to. In this Appendix, I give a standard derivation of this theory starting from GR and the Einstein–Hilbert action. The same theory can be deduced from the axioms of relational field theory without presupposing spacetime and without starting from the Einstein–Hilbert action, as was done in Chapter 7. This derivation assumes less and should be considered more fundamental than that of

Arnowitt, Deser and Misner. However, I felt that the junior readers should be aware of the standard derivation. Finally, Appendix B contains a series of results and derivations that are useful and referenced to throughout the text, but which are moved to the end of the book for the sake of clarity of exposition.

1.2 Notation

In the text we use a notation according to which the Greek indices $\mu, \nu, \ldots$ go from 0 to 3, while the lowercase Latin indices from the middle of the alphabet $i, j, k, \ell, m \ldots$ are spatial and go from 1 to 3. We assume a Lorenzian signature $(-, +, +, +)$. The lowercase Latin indices from the beginning of the alphabet a, b, c, refer to the particle number and go from 1 to N. Three-dimensional vectors will be indicated with Latin or Greek bold letters, $\mathbf{q}, \mathbf{p}, \boldsymbol{\theta}, \boldsymbol{\omega}, \ldots$, while three-dimensional matrices will be uppercase Roman or Greek $\Omega, \Theta, \mathrm{U}, \mathrm{I}, \ldots$. The spatial Laplacian $g_{ij}\nabla^i\nabla^j$ will be indicated with the symbol Δ, while for the d'Alembertian $g_{\mu\nu}\nabla^\mu\nabla^\nu$ I will use the symbol $\Box$. The (spatial) conformal Laplacian $8\Delta - R$ will be indicated with the symbol $\bigcirc$.

1.3 Shape Dynamics in a nutshell

SD is a field theory that describes gravity in a different way to GR. However the differences between the two theories are subtle: in most situations they are indistinguishable.

1.3.1 SD is a gauge theory of spatial conformal (Weyl) symmetry

SD and GR are two different gauge theories defined in the same phase space, both of which admit a particular gauge-fixing in which they coincide. This does not guarantee complete equivalence between the two theories: a gauge-fixing is, in general, not compatible with every solution of a theory, in particular due to global issues. The equivalence between SD and GR therefore fails in some situations.

What distinguishes SD from GR as a fundamental theory of gravity is its different *ontology*.

Firstly, SD does without spacetime: the existence of a pseudo-Riemannian four-dimensional manifold with a Lorentzian signature is not assumed among the axioms of the theory. Instead, the primary entities in SD are three-dimensional geometries that are fitted together by relational principles into a 'stack' whose structural properties can be identified in some, but not all, cases with those of a four-dimensional spacetime which satisfies Einstein's field equations. The closest agreement with GR occurs if the three-geometries are spatially closed when the relational principles of SD are fully implemented. However, there is also interest in the partial implementation of SD's relational principles in cases where the three-geometries are asymptotically flat.

Secondly, the spatial geometries which make the configuration space of SD are not Riemannian. They are *conformal geometries*, defined as equivalence classes of metrics

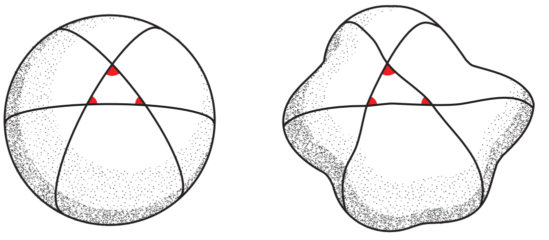

Figure 1 *Conformal transformation of a two-dimensional sphere. The triangle defined by the intersection of the three curves is transformed in such a way that its area and the lengths of its three edges are changed, but the three internal angles (in red) are left invariant.*

under position-dependent conformal transformations (sometimes called 'Weyl trasformations'; the fourth power of ϕ is chosen to simplify the transformation law of the scalar curvature R):

$$\{g_{ij} \sim g'_{ij} \;\; \textit{if} \;\; g'_{ij} = \phi^4 \, g_{ij} \, , \;\; \phi(x) > 0 \; \forall x\} \, . \tag{1}$$

Conformal transformations change lengths and preserve only angles (see Fig. 1). Therefore, a conformal geometry presupposes less than a Riemannian geometry, for which lengths determined by the metric are considered physical. What is physical in SD is the *conformal structure*, which is the *angle-determining* part of the metric. Lengths can be changed arbitrarily and locally by a conformal transformation, which is a gauge transformation for SD.

So, SD assumes less structure than GR, but it is in one sense a minimalistic lifting of assumptions: the next thing in order of simplicity after Riemannian geometry is conformal geometry. Some other approaches to quantum gravity are decidedly more radical as regards to the amount of structure they assume: either much more (*e.g.* string theory) or much less (*e.g.* causal sets).

SD is based on fewer and more basic kinematical first principles than GR:

> *Spatial relationalism:* the positions and sizes of objects are defined relative to each other. This determines what the physical configuration space is (see Sec. 3.2). In field theory this principle translates into conformal and diffeomorphism invariance, and the requirement of a spatially closed manifold.
>
> *Temporal relationalism:* the flow of time is solely due to physical changes (see Sec. 5.2).
>
> The *Mach–Poincaré Principle:* a point and a direction (or tangent vector, in its weak form) in the physical configuration space are sufficient to uniquely specify the solution (see Sec. 3.4).

There is no need for general covariance, the relativity principle, the existence of spacetime or the existence of measuring rods and clocks. These concepts emerge from the solutions of SD as characteristic behaviours or useful approximations. In this sense SD is more fundamental than GR because it achieves the same with less. See Part III for the full construction of SD starting from its three first principles.

A common mistake is to regard SD just as a gauge-fixing of GR. It is easy to see that this is not the case: *there are solutions of SD that are not solutions of GR, and vice-versa.*

A satisfactory understanding of the GR solutions which SD excludes and of the SD solutions which GR excludes is still lacking.

Let us now have a brief look at what exactly SD looks like.

1.3.2 Gauge theories are constrained Hamiltonian systems

SD is more naturally formulated as a gauge theory in the Hamiltonian language. Gauge theories are theories with redundancies: one uses more degrees of freedom than necessary in order to attain a simpler and *local* description. In the Hamiltonian picture, this translates into *nonholonomic constraints:* functions of the canonical variables $\chi = \chi(p, q)$ (with some dependency on the momenta everywhere in phase space) which need to vanish on the solutions of the theory $\chi(p,q) \approx 0$.[1] A single constraint identifies a codimension-1 hypersurface in phase space, the *constraint surface*, on which the solutions of the theory are localized. For example, if a gauge constraint can be written as $\chi = p_1$, where p_1 is one of the canonical momenta (as is always possible, thanks to Darboux's theorem [1]) the constraint surface is the hyperplane $p_1 \approx 0$ shown in Fig. 2. However, p_1 also plays the role of the *generator* of gauge transformations, which happen to be the translations in the q_1 direction: through the Poisson bracket it defines a vector field on phase space $\{p_1, \cdot\} = \frac{\partial}{\partial q_1}$, which is parallel to the q_1 axis (see Fig. 2). This vector field generates infinitesimal transformations on phase space (translations in the q_1 direction), and its integral curves are the *gauge orbits* of the transformations. All the points on these curves are gauge-equivalent (they are related by gauge transformations: they have different representations but the same physical content). Moreover, the vector field $\frac{\partial}{\partial q_1}$ is parallel to the constraint surface $p_1 \approx 0$ by construction, and its integral curves

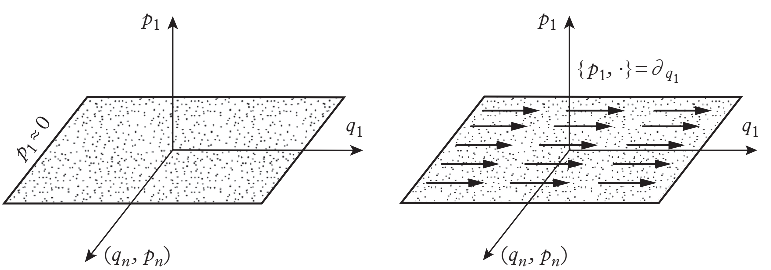

Figure 2 *The constraint surface of a gauge constraint $\chi = p_1$ is represented in phase space, where I put p_1 and q_1 on two axes, and all the other phase-space variables (q_n, p_n), $n = 2, 3\ldots$ are represented collectively on the third axis. On the right, are the vector field generated by p_1 through Poisson-brackets, $\{p_1, \cdot\} = \frac{\partial}{\partial q_1}$, which points towards the q_1 direction. The vector field is parallel to the constraint surface, and its integral curves (the gauge orbits) lie on it.*

[1] With '$\approx$' we mean that the equation holds on the solutions of the constraint equations, following Dirac's notation.

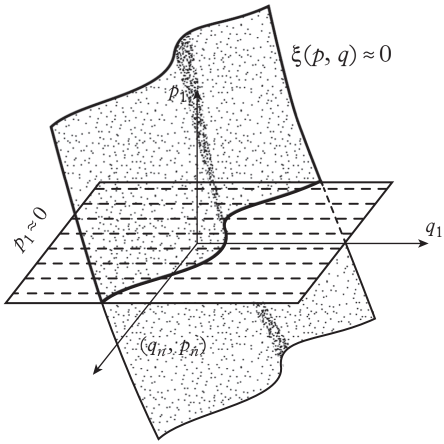

Figure 3 *The concept of gauge-fixing surface: the variable q_1 is unphysical and its value can be taken arbitrarily, therefore we might choose a conventional value for q_1, to be determined by the value of all the other phase-space variables, $q_1 = q_1(q_2, p_2, \ldots)$. One way of obtaining this is to intersect the constraint surface $p_1 \approx 0$ with another surface, $\xi(q, p) \approx 0$, such that it is never parallel to $p_1 \approx 0$ or, at the intersection, 'runs along' the gauge orbits (represented by dashed lines on the constraint surface).*

lie on it. The physical meaning of a gauge constraint $\chi = p_1$ is that the q_1 coordinate is unphysical, like the non-gauge-invariant part of the electromagnetic potentials $\mathbf{A}$ and φ, or like the coordinates of the centre-of-mass of the whole universe.

Since the q_1 coordinate is not physical, we can assign it any value along the solution without changing anything physical. It is often useful (and necessary in quantum mechanics) to fix the value of q_1 by some convention. The standard way of doing it is by choosing a *gauge-fixing:* we specify the value of q_1 as a function of the other variables, $q_1 = q_1(q_2, p_2, \ldots)$. This corresponds to intersecting the constraint surface $p_1 \approx 0$ with another surface $\xi(p, q) \approx 0$ that specifies an intersection submanifold $\{p, q \text{ s.t. } \chi \approx 0, \xi \approx 0\}$ (see Fig. 3). The gauge-fixing should specify the gauge without ambiguity: it has to form a proper intersection with $p_1 \approx 0$, and therefore cannot be parallel to it where they intersect. Moreover, at its intersection with the constraint surface $\chi \approx 0$, the gauge-fixing surface $\xi \approx 0$ cannot 'run along' (be tangent to) any of the gauge orbits: in that case there would be more than one value of q_1 that would correspond to the same value of $q_2, p_2, \ldots$. These two conditions define a good gauge-fixing surface. For details on constrained Hamiltonian systems and gauge theories, see Sec. 6.2.

1.3.3 GR as a constrained Hamiltonian theory

Arnowitt, Deser and Misner (ADM) formulated GR in the Hamiltonian language. They foliated spacetime into a stack of spatial hypersurfaces and split the 4-metric $g_{\mu\nu}$ into a spatial part g_{ij} and four additional components g_{0i} and g_{00}. The spatial metric components g_{ij} represent the canonical variables, and their momenta p^{ij} are related to the extrinsic curvature of the spatial hypersurface with respect to its embedding in spacetime. The g_{0i} and g_{00} components (or better some combinations thereof) enter the action without time derivatives, and are therefore Lagrange multipliers. They are associated with four *local* constraints (meaning one constraint per spatial point). These constraints are the so-called 'superhamiltonian' $\mathcal{H}$ and 'supermomentum' $\mathcal{H}^i$ constraints. Here I will call them the 'Hamiltonian' and the 'diffeomorphism' constraints. The diffeomorphism constraint admits a simple geometrical interpretation: its vector flow sends configuration variables into themselves (one says it generates 'point transformations' $\mathcal{H}^i : g_{ij} \to g_{ij}$), and there is no doubt about its being a gauge constraint.

For the Hamiltonian constraint things are not that simple: it is quadratic in the momenta, and its vector flow does not admit the interpretation of a point transformation (it sends g_{ij}s into both g_{ij}s and p^{ij}s). There is a large literature on the problem of interpreting $\mathcal{H}$. If it is interpreted as a gauge constraint, one would end up with the paradoxical conclusion that the dynamical evolution of GR is just a gauge transformation. There are also huge problems with the definition of what people call 'Dirac observables': quantities whose Poisson brackets with all the first-class constraints vanish on the constraint surface (meaning they must be invariant under the associated gauge transformations). In GR's case, that definition would lead to observables which are constants of motion and do not evolve ('perennials', as Kuchař called them [2]). Kuchař advocated a different notion of observables, namely an idea that they are only required to be invariant under diffeomorphisms. These would evolve, but there are too many: they would depend on *three* polarizations of gravitational waves, while it is widely agreed that gravitational waves have *two* physical polarizations.

The fact that $\mathcal{H}$ is quadratic in the momenta also causes major problems in its quantization. It leads to the notorious 'Wheeler–DeWitt equation', for which there are many unsolved difficulties, above all its 'timelsss' nature, but also ordering ambiguities and coincidence limits. The ADM formulation of GR is detailed in Appendix A, and the problems with this theory which lead to the introduction of SD are explained at the end of Chapter 7 and in Chapter 8.

As illustrated in Fig. 4, SD is based on the identification of the part of $\mathcal{H}$ which is not associated with a gauge redundancy and takes it as the generator of the dynamics. The rest of $\mathcal{H}$ is interpreted as a gauge-fixing for another constraint, $\mathcal{C}$. This constraint is linear in momenta and generates genuine gauge transformations, constraining the physical degrees of freedom to be two per point.

1.3.4 Not every constraint corresponds to gauge redundancy

That this is the case is pretty obvious: think about a particle constrained on a sphere or a plane, *i.e.* a holonomic constraint. Such a constraint obviously has nothing to do with gauge redundancy. However, there are constraints which Dirac [3, 4] argued can always be related to gauge symmetries: they are the so-called 'first-class' constraints. Being first-class means that they close an algebra under Poisson brackets with each other and with the Hamiltonian of the system. If that is the case, Dirac showed that one has freely specifiable variables in the system, one for each first-class constraint, and changing these variables does not change the solutions of the theory. But Barbour and Foster [5] have pointed out that the premises under which Dirac obtained his result do not hold in the important case in which the canonical Hamiltonian vanishes. In that case the Hamiltonian is just a linear combination of constraints, but that does not prevent the theory from having sensible solutions. The solutions will be curves in phase space,

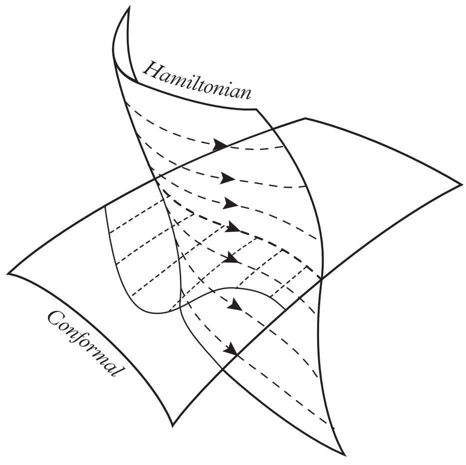

Figure 4 *A schematic representation of the phase space of GR. In it, two constraints coexist, which are good gauge-fixings for each other and are both first-class with respect to the diffeomorphism constraint. One is the Hamiltonian constraint and the other is the conformal (Weyl) constraint. The Hamiltonian constraint is completely gauge-fixed by the conformal constraint except for a single residual global constraint. It Poisson commutes with the conformal constraint and generates a vector flow on the Hamiltonian constraint surface (represented in the figure), which is parallel to the conformal constraint surface. This vector flow generates the time evolution of the system in the intersection between the two surfaces. Any solution can then be represented in an arbitrary conformal gauge by lifting it from the intersection to an arbitrary curve on the conformal constraint surface. All such lifted curves are gauge-equivalent solutions of a conformal gauge theory with a conformally invariant Hamiltonian.*

and will still possess one freely specifiable variable for each constraint—however, one of these redundancies will not change the curve in phase space: it will just change its *parametrization*. Therefore one of the first-class constraints of the system will not be related to any gauge redundancy: there is not an associated unphysical 'q_1' direction as in the example above. This counterexample to Dirac's statement is very important because it is realized in the theory we care about the most: General Relativity. One of the (many) constraints of GR *should not* be associated with gauge redundancy. The Barbour–Foster argument is explained at the end of Sec. 6.2.

1.3.5 SD reinterprets $\mathcal{H}$ as a gauge-fixing of conformal symmetry

SD identifies another constraint surface $\mathcal{C} \approx 0$ in the phase space of GR, which is a good gauge-fixing for the Hamiltonian constraint. This gauge-fixing though happens to be also a gauge symmetry generator. It generates conformal transformations (1) of the spatial metric, with the additional condition that these transformations must preserve the total volume of space $V = \int \mathrm{d}^3x\sqrt{g}$. The constraint $\mathcal{C}$, in addition, happens to close a first-class system with the diffeomorphism constraint $\mathcal{H}^i$, therefore it is a matter of opinion whether it is $\mathcal{C}$ that gauge-fixes the system $(\mathcal{H}, \mathcal{H}^i)$ or it is $\mathcal{H}$ which gauge-fixes

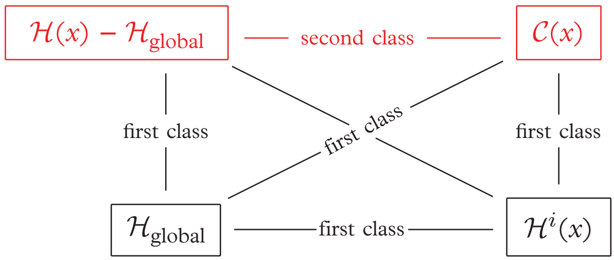

Figure 5 *Scheme of the constraints of GR and of SD. GR's Hamiltonian constraint $\mathcal{H}$ has been split into the global part $\mathcal{H}_{global}$ which is first class with respect to the conformal constraint $\mathcal{C}$ and the part that is purely second class, $\mathcal{H}(x) - \mathcal{H}_{global}$. This second class system admits two first class subalgebras: the lower-left triangle, which constitutes the constraint algebra of GR; and the lower-right triangle, making the constraint algebra of SD.*

$(\mathcal{C}, \mathcal{H}^i)$. If the real physics only lies in the intersection between $\mathcal{C} \approx 0$ and $\mathcal{H} \approx 0$ (which is the big assumption at the basis of SD, and does not hold if spacetime is assumed as an axiom), then the logic can be reversed and the Hamiltonian constraint can be interpreted as a special gauge-fixing for the conformal constraint, see Fig. 5. Then gravity can be reinterpreted as a gauge theory of conformal transformations, which admits a gauge-fixing that is singled out by some special properties. These properties, as I will show, have to do with the fact that it gives a 'natural' notion of scale and proper time, which agree (most of the time) with those measured by physical rods and clocks.

1.3.6 SD's Hamiltonian constraint

$\mathcal{H}$ and $\mathcal{C}$ do not entirely gauge-fix each other: there is a single linear combination of $\mathcal{H}(x)$ which is first class with respect to $\mathcal{C}$. This linear combination, $\mathcal{H}_{\text{global}} = \int d^3x\, N_{\text{CMC}}(x)\, \mathcal{H}(x)$, is a single global constraint whose vector flow is parallel to both the $\mathcal{C} \approx 0$ and the $\mathcal{H} \approx 0$ surfaces on their intersection. This vector flow generates an evolution in the intersection: it has to be interpreted as the generator of time evolution. It is the part of our constraints which is not associated with a gauge redundancy and is instead associated with time reparametrizations of the solutions of the theory.

1.3.7 The 'Linking Theory'

SD pays a price for its conceptual simplicity: the generator of the evolution $\mathcal{H}_{\text{global}}$ contains the solution to a differential equation, N_{CMC}, and therefore is a nonlocal expression. But one can recover a local treatment by enlarging the phase space. SD can in fact be considered as one of the possible gauge-fixings of a first-class theory which is local (its constraints are local) and lives in a larger phase space than that of GR.

This phase space is obtained from that of GR by adjoining a scalar field ϕ and its conjugate momentum π. The larger theory (called the 'Linking Theory') is defined by the constraints of GR, $\mathcal{H}$ and $\mathcal{H}^i$, but expressed in terms of (volume-preserving) conformally-transformed metrics $e^{4\hat{\phi}} g_{ij}$ and momenta $e^{-4\hat{\phi}} \left[p^{ij} - \frac{1}{3}(1 - e^{6\hat{\phi}}) \sqrt{g} g^{ij} \int p/V \right]$, where $\hat{\phi} = \phi - \frac{1}{6} \log(\int \mathrm{d}^3x \sqrt{g} \exp(6\phi)/V)$ and $V = \int \mathrm{d}^3x \sqrt{g}$. In addition, one has a modified conformal constraint which includes a term that transforms ϕ. The new constraint is $\mathcal{Q} = 4\pi - \mathcal{C}$ and generates simultaneous translations of ϕ and (volume-preserving) conformal transformations of π, so that the combination $e^{4\hat{\phi}} g_{ij}$, is left invariant. The constraint $\mathcal{Q}$ is now first class with respect to $\mathcal{H}$ and $\mathcal{H}^i$. By completely gauge-fixing $\mathcal{Q}$, for example with the condition $\phi \approx 0$, one obtains GR. On the other hand, one can use a different gauge-fixing, namely $\pi \approx 0$, which is first class with respect to $\mathcal{Q}$, but gauge-fixes $\mathcal{H}$ almost entirely, leaving only the global part $\mathcal{H}_{\text{global}}$ untouched.

One can then work with the Linking Theory, where all the equations of motion and constraints are local (apart from their dependence on the total volume), and work out the solutions in this framework. As long as the solution is compatible with the gauge-fixing $\pi \approx 0$, it is a legitimate SD solution.

All the details of the SD construction can be found in Chapter 9 and in Part IV.

1.3.8 The present status of SD

The greatest hope of SD's is to provide a new approach to quantum gravity based on a sum over the histories of conformal 3-geometries. Such an approach is so far unexplored: 3D conformal symmetry plays a role in some quantum gravity proposals, *e.g.* Hořava–Lifshitz gravity or Causal Dynamical Triangulations, in which there is evidence that the theory admits conformally-invariant fixed points. However, such symmetries are only asymptotic and do not characterize the physically relevant regimes of these theories. In particular, it seems that a common feature of quantum gravity models is the generation of additional dynamical degrees of freedom at the quantum level, *e.g.* the scalar mode in Hořava–Lifshitz gravity. A quantum gravity theory compatible with the principles of SD should, presumably, keep only two dynamical degrees of freedom all the way through its renormalization group orbit. Such a proposal seems to contradict the fact that quantum mechanics introduces a scale in physics through $\hbar$, and therefore breaks conformal invariance. This, however, should not be taken as a sacred fact of nature: the fact that $\hbar$ is dimensionful can be a consequence of the fact that we measure its effects in a subsystem of the universe, and its scale might admit a relational expression in terms of the state of the matter in the universe. Indeed, in [6] we formulated a toy model of scale-invariant non-relativistic quantum mechanics, in which the fundamental ontology is that of a wavefunction evolving on shape space. The role of $\hbar$ is played by a dimensionless quantity which is related to the 'spreading' of the wavefunction on shape space. Such a quantity is intimately related to the particle equivalent of York time, the dilatational momentum D (in appropriate units $\hbar$ and D have the same dimensions). The $\hbar$ which we use to describe phenomena we observe in the laboratory emerges, in such a model, when we concentrate on subsystems of the universe and model them semiclassically as quantum fluctuations around a classical solution. In this framework, doubling the size of the universe and simultaneously doubling $\hbar$ has no effect.

The chance of exploring an uncharted road to quantum gravity is one of the main motivations behind SD. However, so far, the work of researchers in the field has been (mostly) limited to the classical theory. The reason for this is that before going quantum, we need to learn from scratch how to do physics without spacetime and relying only on the 3D conformal geometry of space.

The most important conceptual point that should be clarified at the classical level is whether SD is equivalent or not to GR. This should be investigated in those situations in which GR develops singularities, or when spacetime ceases to be CMC-foliable. The first evidence that SD can do better than GR in a singular situation was [7], in which Budd and Koslowski studied homogeneous cosmological solutions in two spatial dimensions with the topology of a torus. In this case the dynamics is that of the 'Bianchi I' model, and inevitably it reaches 'crushing' singularities in which $\det g \to 0$. In such situations the spacetime description ceases to make sense. However one can evolve the conformal geometry of space through those singularities, simply by requiring continuity of the shape degrees of freedom. This result could have simply been a fluke of a lower-dimensional toy model, however in the recent [8] we studied the much more complicated 3D 'Bianchi IX' model of homogeneous cosmology (this model is described in detail in Sec. 13.1). In [8] we showed that at the singularity it is only the spatial volume and its conjugate momentum, the York time, which are singular. The conformal geometry at the singularity is degenerate, because it is flattened to a two-dimensional object, but the shape degrees of freedom are not singular. The situation is analogous to that of a 3-body model in which the three particles go collinear: the triangle they describe is degenerate and its area is zero, but as shapes, the collinear configurations are perfectly regular. The shape of freedom can be continued through the singularity in a unique way just by requiring continuity, and on the other side the dynamics continues undisturbed, following the Bianchi IX equations of motion. In this way we end up joining two cosmological solutions of GR at the singularity: each half is an acceptable solution of GR and can be described as a spacetime, but the whole solution cannot. In this sense we proved that SD admits solutions which are not in GR, and it can do better than GR in dealing with singularities.

Another situation in which GR predicts singularities is the case of black holes. Moreover, it is known that CMC foliations have a *singularity-avoiding* property [9, 10] in Schwarzschild's spacetime, so one could legitimately conjecture that the shape-dynamical description of black holes may be different from that of GR. The first study on the subject was Gomes' paper [11] studying asymptotically flat, spherically symmetric vacuum solutions of ADM gravity in *maximal slicing*. This solution is derived and discussed in Sec. 13.4. Interestingly, this solution does not have any singularity: it consists of a 'wormhole' geometry with two asymptotically flat ends, and a 'throat', that is, a sphere of minimal area. The geometry cannot support any concentric sphere (*i.e.* spheres centred around the centre of symmetry) of area smaller than the throat. Of course, I am now talking about the Riemannian geometry of spatial slices, which is not an observable property of SD: all of these geometries are conformally related to the