INTERNATIONAL SERIES OF MONOGRAPHS ON PHYSICS • 155

Quantum Gravity

Third Edition

CLAUS KIEFER

OXFORD SCIENCE PUBLICATIONS

International Series of Monographs on Physics

155. C. Kiefer: *Quantum gravity, Third edition*
154. L. Mestel: *Stellar magnetism, Second edition*
153. R. A. Klemm: *Layered superconductors, Volume 1*
152. E. L. Wolf: *Principles of electron tunneling spectroscopy, Second edition*
151. R. Blinc: *Advanced ferroelectricity*
150. L. Berthier, G. Biroli, J.-P. Bouchaud, W. van Saarloos, L. Cipelletti: *Dynamical heterogeneities in glasses, colloids and granular media*
149. J. Wesson: *Tokamaks, Fourth edition*
148. H. Asada, T. Futamase, P. Hogan: *Equations of motion in general relativity*
147. A. Yaouanc, P. Dalmas de Réotier: *Muon spin rotation, relaxation, and resonance*
146. B. McCoy: *Advanced statistical mechanics*
145. M. Bordag, G. L. Klimchitskaya, U. Mohideen, V. M. Mostepanenko: *Advances in the Casimir effect*
144. T. R. Field: *Electromagnetic scattering from random media*
143. W. Götze: *Complex dynamics of glass-forming liquids – a mode-coupling theory*
142. V. M. Agranovich: *Excitations in organic solids*
141. W. T. Grandy: *Entropy and the time evolution of macroscopic systems*
140. M. Alcubierre: *Introduction to 3+1 numerical relativity*
139. A. L. Ivanov, S. G. Tikhodeev: *Problems of condensed matter physics – quantum coherence phenomena in electron–hole and coupled matter–light systems*
138. I. M. Vardavas, F. W. Taylor: *Radiation and climate*
137. A. F. Borghesani: *Ions and electrons in liquid helium*
135. V. Fortov, I. Iakubov, A. Khrapak: *Physics of strongly coupled plasma*
134. G. Fredrickson: *The equilibrium theory of inhomogeneous polymers*
133. H. Suhl: *Relaxation processes in micromagnetics*
132. J. Terning: *Modern supersymmetry*
131. M. Mariño: *Chern–Simons theory, matrix models, and topological strings*
130. V. Gantmakher: *Electrons and disorder in solids*
129. W. Barford: *Electronic and optical properties of conjugated polymers*
128. R. E. Raab, O. L. de Lange: *Multipole theory in electromagnetism*
127. A. Larkin, A. Varlamov: *Theory of fluctuations in superconductors*
126. P. Goldbart, N. Goldenfeld, D. Sherrington: *Stealing the gold*
125. S. Atzeni, J. Meyer-ter-Vehn: *The physics of inertial fusion*
123. T. Fujimoto: *Plasma spectroscopy*
122. K. Fujikawa, H. Suzuki: *Path integrals and quantum anomalies*
121. T. Giamarchi: *Quantum physics in one dimension*
120. M. Warner, E. Terentjev: *Liquid crystal elastomers*
119. L. Jacak, P. Sitko, K. Wieczorek, A. Wojs: *Quantum Hall systems*
117. G. Volovik: *The Universe in a helium droplet*
116. L. Pitaevskii, S. Stringari: *Bose–Einstein condensation*
115. G. Dissertori, I. G. Knowles, M. Schmelling: *Quantum chromodynamics*
114. B. DeWitt: *The global approach to quantum field theory*
113. J. Zinn-Justin: *Quantum field theory and critical phenomena, Fourth edition*
112. R. M. Mazo: *Brownian motion – fluctuations, dynamics, and applications*
111. H. Nishimori: *Statistical physics of spin glasses and information processing – an introduction*
110. N. B. Kopnin: *Theory of nonequilibrium superconductivity*
109. A. Aharoni: *Introduction to the theory of ferromagnetism, Second edition*
108. R. Dobbs: *Helium three*
107. R. Wigmans: *Calorimetry*
106. J. Kübler: *Theory of itinerant electron magnetism*
105. Y. Kuramoto, Y. Kitaoka: *Dynamics of heavy electrons*
104. D. Bardin, G. Passarino: *The Standard Model in the making*
103. G. C. Branco, L. Lavoura, J. P. Silva: *CP Violation*
102. T. C. Choy: *Effective medium theory*
101. H. Araki: *Mathematical theory of quantum fields*
100. L. M. Pismen: *Vortices in nonlinear fields*
98. K. H. Bennemann: *Nonlinear optics in metals*
94. S. Chikazumi: *Physics of ferromagnetism*
91. R. A. Bertlmann: *Anomalies in quantum field theory*
90. P. K. Gosh: *Ion traps*
87. P. S. Joshi: *Global aspects in gravitation and cosmology*
86. E. R. Pike, S. Sarkar: *The quantum theory of radiation*
83. P. G. de Gennes, J. Prost: *The physics of liquid crystals*
73. M. Doi, S. F. Edwards: *The theory of polymer dynamics*
69. S. Chandrasekhar: *The mathematical theory of black holes*
51. C. Møller: *The theory of relativity*
46. H. E. Stanley: *Introduction to phase transitions and critical phenomena*
32. A. Abragam: *Principles of nuclear magnetism*
27. P. A. M. Dirac: *Principles of quantum mechanics*
23. R. E. Peierls: *Quantum theory of solids*

Quantum Gravity

Third Edition

Claus Kiefer
Institute for Theoretical Physics
University of Cologne

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.
If 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

Third Edition published in 2012
Impression: 1

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Library of Congress Control Number: 2012932071

ISBN 978–0–19–958520–5

Printed in Great Britain by
CPI Group (UK) Ltd, Croydon CR0 4YY

Preface to the Third Edition

In my third edition, I have revised the text in many places throughout the book. On the one hand, I have spent considerable effort on improving its conceptual structure and its pedagogical aspects. On the other hand, I have updated and extended the presentation of the topics discussed in the earlier editions and added new material on topics such as primordial black holes, black holes in accelerators, singularity avoidance, quantum-gravity phenomenology, string field theory, and Hořava–Lifshits gravity. Whenever possible, I have shown the links between the various approaches to quantum gravity. As in the previous editions, the emphasis was on writing an introductory and coherent text and not an encyclopedia of all existing approaches.

I thank all the readers who have commented on various parts of the earlier editions or who have pointed out the occurrence of typos. Comments would be greatly appreciated also for the third edition. They can be sent to my e-mail address `kiefer@thp.uni-koeln.de`.

For helpful discussions and critical comments while preparing the third edition, I owe my thanks to Julian Barbour, Andrei Barvinsky, Martin Bojowald, Bianca Dittrich, Domenico Giulini, Herbert Hamber, Friedrich Hehl, Manuel Krämer, Claus Lämmerzahl, Renate Loll, Paulo Vargas Moniz, Hermann Nicolai, Marcel Reginatto, Martin Reuter, Mark Roberts, Christian Steinwachs, Kurt Sundermeyer, Stefan Theisen, and H.-Dieter Zeh. I am grateful in particular to Julian Barbour for going through the whole second edition with me and making numerous invaluable suggestions for improving style and conceptual clarity. I also thank the Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Potsdam, for its kind hospitality while part of the work on the new edition was done. I have also profited much from discussions on historical aspects of quantum-gravity research at the Max Planck Institute for the History of Science, Berlin, and I thank Jürgen Renn for giving me this opportunity.

As before, I want to thank Oxford University Press for their most efficient co-operation.

Cologne
January 2012

Claus Kiefer

This page intentionally left blank

Preface to the Second Edition

The construction of a consistent theory of quantum gravity continues to be the major open problem in fundamental physics. The present second edition of my book is an extended and revised version of the first edition. It contains, in particular, new sections on asymptotic safety, dynamical triangulation, primordial black holes, the information-loss problem, loop quantum cosmology, and other topics. The text has been revised throughout.

I thank all the readers who have commented on various parts of the first edition or who have pointed out the occurrence of typos. Comments would be greatly appreciated also for the second edition. They can be sent to my e-mail address `kiefer@thp.uni-koeln.de`.

For helpful discussions and critical comments while preparing the second edition I am grateful to Mark Albers, Andrei Barvinsky, Martin Bojowald, Friedrich Hehl, Gerhard Kolland, Renate Loll, Hermann Nicolai, Martin Reuter, Barbara Sandhöfer, Stefan Theisen, and H.-Dieter Zeh.

Last but not least I want to thank Oxford University Press, and in particular Sonke Adlung, for their efficient cooperation.

Cologne Claus Kiefer
December 2006

This page intentionally left blank

Preface

The unification of quantum theory with Einstein's theory of general relativity is perhaps *the* biggest open problem of theoretical physics. Such a theory is not only needed for conceptual reasons, but also for the understanding of fundamental issues such as the origin of the Universe, the final evaporation of black holes, and the structure of space and time.

Historically, the oldest approach is the direct quantization of Einstein's theory of general relativity, an approach which is still being actively pursued. This includes covariant methods such as path-integral quantization as well as canonical methods like the Wheeler–DeWitt approach or the more recent loop quantization. Although one arrives at a perturbatively non-renormalizable theory, quantum general relativity can yield physically interesting results in both the perturbative and the non-perturbative regimes. It casts light, in particular, on the fundamental nature of space and time.

The second main approach is string theory. It encapsulates the idea that the problem of constructing a viable quantum theory of gravity can only be solved within a unification of all interactions. In this respect, it goes far beyond quantum general relativity. From a methodological point of view, however, string theory does not stand much apart from it. It is a natural extension of perturbative quantum gravity (from which it inherits the concept of a graviton), and methods of constrained quantization, which are crucial for canonical quantum gravity, appear at key stages in the theory.

Whereas there exist excellent textbooks that discuss string theory at great depth, the present monograph is the first that, to my knowledge, covers quantum gravity in this broad sense. The main part of the book is devoted to general concepts, the quantization of general relativity, and applications to cosmology and black holes. String theory is discussed from the point of view of its quantum gravitational aspects and its connection to other approaches. The edifice of theoretical physics cannot be completed without the conceptual unification that will be provided by quantum gravity. I hope that my book will convince my readers of this outstanding problem and encourage them to work on its solution.

This book has grown out of lectures that I gave at the Universities of Zürich, Freiburg, and Cologne between 1990 and 2003. My main intention is to discuss the general features that a quantum theory of gravity is expected to show and to give an up-to-date overview of the main approaches. The reader is assumed to have some familiarity with general relativity and quantum field theory. Comments can be sent to my e-mail address `kiefer@thp.uni-koeln.de` and are highly welcome.

It is clear that my book could not have been written in this form without the influence of many people over the past 20 years. I am in particular indebted to H.-Dieter Zeh for encouraging me to enter this field of research and for many stimulating and inspiring interactions. I also thank Norbert Straumann in Zürich and Hartmann Römer in Freiburg for providing me with the excellent working conditions that gave me the freedom to follow the research I wanted to do. Many people have read through drafts of this book, and their critical comments have helped me to improve various parts. I owe thanks to Julian Barbour, Andrei Barvinsky, Domenico Giulini, Alexander Kamenshchik, Thomas Mohaupt, Paulo Moniz, Andreas Rathke, Thomas Thiemann, and H.-Dieter Zeh. I am also deeply indebted to them and also, for discussions over many years, to Mariusz Dąbrowski, Lajos Diósi, Petr Hájíček, Erich Joos, Jorma Louko, David Polarski, T. P. Singh, Alexei Starobinsky, and Andreas Wipf. I have actively collaborated with most of the aforementioned theoreticians and I want to take this opportunity to thank them for the pleasure that I could experience during these collaborations.

Cologne
January 2004

Claus Kiefer

Contents

1	**Why quantum gravity?**	**1**
1.1	Quantum theory and the gravitational field	1
1.2	Problems of a fundamentally semiclassical theory	15
1.3	Approaches to quantum gravity	24
2	**Covariant approaches to quantum gravity**	**28**
2.1	The concept of a graviton	28
2.2	Path-integral quantization	42
2.3	Quantum supergravity	74
3	**Parametrized and relational systems**	**78**
3.1	Particle systems	78
3.2	The free bosonic string	84
3.3	Parametrized field theories	90
3.4	Relational dynamical systems	96
3.5	General remarks on constrained systems	100
4	**Hamiltonian formulation of general relativity**	**106**
4.1	The seventh route to geometrodynamics	106
4.2	The 3+1 decomposition of general relativity	113
4.3	Canonical gravity with connections and loops	132
5	**Quantum geometrodynamics**	**141**
5.1	The programme of canonical quantization	141
5.2	The problem of time	144
5.3	The geometrodynamical wave function	152
5.4	The semiclassical approximation	172
6	**Quantum gravity with connections and loops**	**190**
6.1	Connection and loop variables	190
6.2	Quantization of area	198
6.3	Quantum Hamiltonian constraint	203
7	**Quantization of black holes**	**208**
7.1	Black-hole thermodynamics and Hawking radiation	208
7.2	Canonical quantization of the Schwarzschild black hole	221
7.3	Black-hole spectroscopy and entropy	228
7.4	Quantum theory of collapsing dust shells	233
7.5	The Lemaître–Tolman–Bondi model	241
7.6	The information-loss problem	249
7.7	Primordial black holes	252

8 Quantum cosmology 260
8.1 Minisuperspace models 260
8.2 Introduction of inhomogeneities 275
8.3 Boundary conditions 279
8.4 Loop quantum cosmology 289
8.5 On singularity avoidance 294

9 String theory 298
9.1 General introduction 298
9.2 Quantum-gravitational aspects 303
9.3 String field theory 327

10 Phenomenology, decoherence, and the arrow of time 329
10.1 Quantum-gravity phenomenology 329
10.2 Decoherence and the quantum universe 331
10.3 Arrow of time 343
10.4 Reflections and outlook 349

References 354

Index 389

1

Why quantum gravity?

1.1 Quantum theory and the gravitational field

1.1.1 Introduction

Quantum theory seems to be a universal theory of nature. More precisely, it provides a general framework for all theories describing particular interactions. Quantum theory has passed a plethora of experimental tests and is considered a well-established theory, except for the ongoing discussion about its interpretational foundations.

The only interaction that has not been fully accommodated within quantum theory is the gravitational field, the oldest known interaction. It is described very successfully by a classical (i.e. non-quantum) theory, Einstein's *general theory of relativity* (GR), also called *geometrodynamics*. From a theoretical, or even aesthetic, point of view, it is highly appealing, since the fundamental equations can be formulated in simple geometrical terms. Moreover, there now exist plenty of experimental tests that have been passed by this theory without problems. One particularly impressive example is the case of the binary pulsar PSR 1913+16: the decrease of its orbital period can be fully explained by the emission of gravitational waves as predicted by GR. The accuracy of this test is only limited by the accuracy of clocks on Earth, which according to recent proposals for rubidium fountain clocks (Fertig and Gibble 2000) should approach an accuracy of about 10^{-16} (such a clock would go wrong by less than 1 s during a time as long as the age of the Universe). The precision is so high that one even needs to model the gravitational influence of the Milky Way on the binary pulsar in order to find agreement with the theoretical prediction (Damour and Taylor 1991). There exist phenomena that could point to a more fundamental theory than GR (dark matter, dark energy, and the Pioneer anomaly), but this is not yet clear.

The formalism of general relativity is discussed in many textbooks; see, for example, Hawking and Ellis (1973), Misner *et al.* (1973), Straumann (2004), or Wald (1984). It can be defined by the Einstein–Hilbert action,

$$S_{\mathrm{EH}} = \frac{c^4}{16\pi G}\int_{\mathcal{M}} \mathrm{d}^4x\,\sqrt{-g}\,(R-2\Lambda) - \frac{c^4}{8\pi G}\int_{\partial\mathcal{M}} \mathrm{d}^3x\,\sqrt{h}K. \tag{1.1}$$

Note that $c^4/16\pi G \approx 2.41\times 10^{47}\ \mathrm{g\,cm\,s^{-2}} \approx 2.29\times 10^{74}\ (\mathrm{cm\,s})^{-1}\ \hbar$. The integration in the first integral of (1.1) covers a region $\mathcal{M}$ of the space–time manifold, and the second integral is defined on its boundary $\partial\mathcal{M}$, which is assumed to be space-like. The integrand of the latter contains the determinant, h, of the three-dimensional metric on the boundary, and K is the trace of the second fundamental form (see Section 4.2.1).

That a surface term is needed in order to obtain a consistent variational principle had already been noted by Einstein (1916*a*).

In addition to the action (1.1), one considers actions for non-gravitational fields, in the following called $S_{\rm m}$ ('matter action'). They give rise to the energy–momentum tensor

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta S_{\rm m}}{\delta g^{\mu\nu}}, \tag{1.2}$$

which acts as a 'source' of the gravitational field. In general, it does not coincide with the canonical energy–momentum tensor. From the variation of $S_{\rm EH} + S_{\rm m}$, the Einstein field equations are obtained,

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu} - \Lambda g_{\mu\nu}. \tag{1.3}$$

(Our convention is the one used by Misner *et al.* (1973), with a space–time metric of signature $(-1, 1, 1, 1)$, a Riemann tensor defined by $+R^{\alpha}{}_{\beta\gamma\delta} = \Gamma^{\alpha}{}_{\beta\delta,\gamma} - \cdots$, and a Ricci tensor defined by $+R_{\alpha\beta} = R^{\mu}{}_{\alpha\mu\beta}$.)

A natural generalization of general relativity is the Einstein–Cartan theory; see, for example, Hehl (1985) and Hehl *et al.* (1976) for details. It is found by introducing potentials connected with translations[1] and Lorentz transformations (i.e. with the Poincaré group): the tetrad e_n^{μ} for the translations and the connection ω_{μ}^{mk} for the Lorentz transformations. The corresponding gravitational field strengths are the torsion and curvature. Torsion vanishes outside matter and does not propagate, but it is straightforward to formulate extensions with a propagating torsion by gauging the translations and the Lorentz transformations. This then leads to the Poincaré gauge theory; see Gronwald and Hehl (1996). The occurrence of torsion is a natural consequence of the presence of spin currents. Its effects are tiny on macroscopic scales (which is why it has not been seen experimentally), but it should be of high relevance in the microscopic realm, for example, on the scale of the electronic Compton wavelength and in the very early Universe. In fact, the Einstein–Cartan theory is naturally embedded in theories of supergravity (see Section 2.3), where a spin-3/2 particle (the 'gravitino') plays a central role.

In Chapter 2, we shall discuss some 'uniqueness theorems', which state that every theory of the gravitational field must contain GR (or the Einstein–Cartan theory) in an appropriate limit. Generalizations of GR such as the Jordan–Brans–Dicke theory, which contains an additional scalar field in the gravitational sector, are therefore mainly of interest as effective theories arising from fundamental theories such as string theory (see Chapter 9). They are usually not meant as classical alternatives to GR, except for the parametrization of experimental tests. That GR cannot be true at the most fundamental level is clear from the *singularity theorems* (cf. Hawking and Penrose 1996): under very general conditions, singularities in space–time are unavoidable, signalling the breakdown of GR.

[1] "...gravity is that field which corresponds to a gauge invariance with respect to displacement transformations." (Feynman *et al.* (1995, p. 115))

The theme of this book is to investigate the possibilities of unifying the gravitational field with the quantum framework in a consistent way. This may lead to a general avoidance of space–time singularities.

1.1.2 Main motivations for quantizing gravity

The first motivation is **unification**. The history of science shows that a reductionist viewpoint has been very fruitful in physics (Weinberg 1993). The Standard Model of particle physics is a *quantum* field theory that has united in a certain sense all non-gravitational interactions. It has been very successful experimentally, but one should be aware that its concepts are poorly understood beyond the perturbative level; in this sense, the classical theory of GR is in a much better condition.

The universal coupling of gravity to all forms of energy would make it plausible that gravity has to be implemented in a quantum framework too. Moreover, attempts to construct an exact semiclassical theory, where gravity stays classical but all other fields are quantum, have failed up to now; see Section 1.2. This demonstrates, in particular, that classical and quantum *concepts* (phase space versus Hilbert space, etc.) are most likely incompatible.

Physicists have also entertained the hope that unification entails a solution to the notorious divergence problem of quantum field theory; as is shown in Chapter 2, perturbative quantum GR leads to even worse divergences, due to its non-renormalizability, but a full non-perturbative framework without any divergences may exist. In fact, some of the approaches presented in this volume, such as canonical quantum gravity or string theory, are candidates for a divergence-free theory.

The second motivation comes from **cosmology** and **black holes**. As the singularity theorems and the ensuing breakdown of GR demonstrate, a fundamental understanding of the early Universe—in particular, its initial conditions near the 'big bang'—and of the final stages of black-hole evolution requires an encompassing theory. From the historical analogy of quantum mechanics (which—due to its stationary states—has rescued atoms from collapse), the general expectation is that this encompassing theory is a *quantum* theory. Classically, the generic behaviour of a solution to Einstein's equations near a big-bang singularity is assumed to consist of 'BKL oscillations'; see Belinskii *et al.* (1982) and the references therein. A key feature of this scenario is the decoupling of different spatial points. A central demand on a quantum theory of gravity is to provide a consistent quantum description of BKL oscillations.

The concept of an 'inflationary universe'[2] is often invoked to claim that the present universe can have emerged from generic initial conditions. This is only partly true, since one can of course trace back *any* present conditions to the past to find the 'correct' initial conditions. In fact, the crucial point lies in the assumptions that enter the *no-hair conjecture*; see, for example, Frieman *et al.* (1997). This conjecture states that space–time approaches locally a de Sitter form for large times if a (probably effective) cosmological constant is present. The conjecture can be proved, provided some assumptions are made. In particular, it must be assumed that modes on very

[2]Following Harrison (2000), we shall write 'universe' instead of 'Universe' to emphasize that we talk about a *model* of the Universe, in contrast to 'Universe', which refers to 'everything'.

small scales (smaller than the Planck length: see below) are not amplified to cosmological scales. This assumption thus refers to the unknown regime of quantum gravity. Moreover, it seems that the singularity theorems apply even to inflationary cosmology (Borde *et al.* 2003).

It must be emphasized that *if* gravity is quantized, the kinematical non-separability of quantum theory demands that the whole universe must be described in quantum terms. This leads to the concepts of quantum cosmology and the wave function of the universe; see Chapters 8 and 10.

A third motivation is the **problem of time**. Quantum theory and GR (in fact, every general covariant theory) contain drastically different concepts of time (and space–time). Strictly speaking, they are incompatible. In quantum theory, time is an external (absolute) element, *not* described by an operator (in special relativistic quantum field theory, the role of time is played by the external Minkowski space–time). In contrast, in GR, space–time is a dynamical (non-absolute) object.[3] It is clear that a unification of quantum theory with GR must lead to modifications of the concept of time. One might expect that the metric has to be turned into an operator. In fact, as a detailed analysis will show (Chapters 5 and 9), this leads to novel features. Related problems concern the role of background structures in quantum gravity, the role of the diffeomorphism group (Poincaré invariance, as used in ordinary quantum field theory, is no longer a symmetry group), and the notion of 'observables'. That a crucial point lies in the presence of a more general invariance group was already noted by Pauli (1955, p. 267)[4]:

> It seems to me ... that it is not so much the linearity or non-linearity which forms the heart of the matter, but the very fact that here a more general group than the Lorentz group is present ... .

A fourth motivation is provided by the **superposition principle** of quantum theory. Since every quantum object carries energy, it generates a gravitational field. A superposition of a quantum object where, for example, it was at two different places would then entail a corresponding superposition of the respective gravitational fields and thus demand a quantum theory of gravity for its description—unless the superposition principle ceases to be valid at that scale. We shall return to this point at the end of Section 1.2.1.

1.1.3 Relevant scales

In a universally valid quantum theory, genuine quantum effects can occur on any scale, while classical properties are an emergent phenomenon only (see Chapter 10). This is a consequence of the superposition principle. Independent of this, there exist scales where quantum effects of a particular interaction should definitely be non-negligible.

It was already noted by Planck (1899) that the fundamental constants of the speed of light (c), gravitational constant (G), and quantum of action ($\hbar$) can be combined

[3] More precisely, as we shall see in Chapter 4, it is the *three-geometry* that is the dynamical object of GR.

[4] 'Es scheint mir ..., daß nicht so sehr die Linearität oder Nichtlinearität Kern der Sache ist, sondern eben der Umstand, daß hier eine allgemeinere Gruppe als die Lorentzgruppe vorhanden ist ... .'

in a unique way to yield units of length, time, and mass. In Planck's honour, they are called the Planck length, $l_{\rm P}$, Planck time, $t_{\rm P}$, and Planck mass, $m_{\rm P}$, respectively. They are given by the expressions

$$l_{\rm P} = \sqrt{\frac{\hbar G}{c^3}} \approx 1.62 \times 10^{-33} \text{ cm}, \tag{1.4}$$

$$t_{\rm P} = \frac{l_{\rm P}}{c} = \sqrt{\frac{\hbar G}{c^5}} \approx 5.39 \times 10^{-44} \text{ s}, \tag{1.5}$$

$$m_{\rm P} = \frac{\hbar}{l_{\rm P} c} = \sqrt{\frac{\hbar c}{G}} \approx 2.18 \times 10^{-5} \text{ g} \approx 1.22 \times 10^{19} \text{ GeV}/c^2. \tag{1.6}$$

It must be emphasized that units of length, time, and mass cannot be formed out of G and c (GR) or out of $\hbar$ and c (quantum theory) alone. Instead of the Planck mass (1.6), one sometimes uses the 'reduced Planck mass' defined by

$$M_{\rm P} := \frac{m_{\rm P}}{\sqrt{8\pi}} \approx 2.43 \times 10^{18} \text{ GeV}/c^2. \tag{1.7}$$

This is motivated by the factor $c^4/16\pi G$ in the Einstein–Hilbert action (1.1).

The Planck mass seems to be a rather large quantity by microscopic standards. One has to keep in mind, however, that this mass (energy) must be concentrated in a region of linear dimension $l_{\rm P}$ in order to allow one to see direct quantum-gravity effects. In fact, the Planck scales are attained for an elementary particle whose (reduced) Compton wavelength is (apart from a factor of 2) equal to its Schwarzschild radius,

$$\frac{\hbar}{m_{\rm P} c} \approx R_{\rm S} \equiv \frac{2Gm_{\rm P}}{c^2},$$

which means that the space–time curvature of such an elementary particle would not be negligible. Sometimes (e.g. in cosmology), one also uses the Planck temperature,

$$T_{\rm P} = \frac{m_{\rm P} c^2}{k_{\rm B}} \approx 1.41 \times 10^{32} \text{ K}, \tag{1.8}$$

and the Planck density,

$$\rho_{\rm P} = \frac{m_{\rm P}}{l_{\rm P}^3} \approx 5 \times 10^{93} \frac{\text{g}}{\text{cm}^3}. \tag{1.9}$$

Recalling the physical dimensions of electric charge in terms of length, time, and mass, one could even define a 'Planck charge' through (cf. Gamow *et al.* 2002)

$$Q_{\rm P} = \sqrt{m_{\rm P} l_{\rm P}} \frac{l_{\rm P}}{t_{\rm P}} = \sqrt{G} m_{\rm P} = \sqrt{\hbar c}, \tag{1.10}$$

which is independent of G. The elementary electric charge is then $e = \sqrt{\alpha} Q_{\rm P} \approx 0.085\, Q_{\rm P}$, where α is the fine-structure constant. For two particles with electric charge $Q_{\rm P}$ and mass $m_{\rm P}$, the Coulombian repulsion would exactly compensate the Newtonian attraction.

It is interesting to observe that Planck had introduced his units one year before he wrote the famous paper containing the quantum of action; see Planck (1899). How had this been possible? The constant $\hbar$ appears in Wien's law, $\hbar\omega_{\max} \approx 2.82 k_{\mathrm{B}} T$, which was phenomenologically known at that time. Planck learnt from this that a new constant of nature is contained in this law, and he called it b. Planck concludes his article by writing[5]:

> These quantities retain their natural meaning as long as the laws of gravitation, of light propagation in vacuum, and the two laws of the theory of heat remain valid; they must therefore, if measured in various ways by all kinds of intelligent beings, always turn out to be the same.

It is interesting that similar units had already been introduced by George Johnstone Stoney (1826–1911) (Stoney 1881). Of course, $\hbar$ was not known at that time, but one could (in principle) get the elementary electric charge e from Avogadro's number L and Faraday's number $F = eL$. With e, G, and c, one can construct the same fundamental units as with $\hbar, G$, and c (since the fine-structure constant is $\alpha = e^2/\hbar c \approx 1/137$); therefore, Stoney's units differ from Planck's units by factors of $\sqrt{\alpha}$.

Quite generally, one can argue that there are three fundamental dimensional quantities (cf. Okun 1992). In a paper submitted in October 1927, Gamow, Ivanenko, and Landau emphasized the important role of the Planck units; see Gamow *et al.* (2002) for an English translation of this paper. These units enjoyed wider publicity only after their discussion in Wheeler (1955).

The Planck length is indeed very small. If one imagines an atom to be the size of the Moon's orbit, l_{P} would only be as small as about a tenth of the size of a nucleus. Still, physicists have for some time entertained the idea that something dramatic happens at the Planck length, from the breakdown of the continuum to the emergence of non-trivial topology ('space–time foam'); see, for example, Misner *et al.* (1973). We shall see in the course of this book how such ideas can be made more precise in quantum gravity. Unified theories may contain an intrinsic length scale from which l_{P} may be deduced. In string theory, for example, this is the string length l_{s}. A generalized uncertainty relation shows that scales smaller than l_{s} have no operational significance; see Chapter 9. We also remark that the Einstein–Hilbert action (1.1) is of order $\hbar$ only for $TL \sim l_{\mathrm{P}} t_{\mathrm{P}}$, where the integration in the action is performed over a space–time region of extension TL^3.

Figure 1.1 presents some of the important structures in our Universe in a mass-versus-length diagram. A central role is played by the 'fine-structure constant of gravity',

$$\alpha_{\mathrm{g}} = \frac{G m_{\mathrm{pr}}^2}{\hbar c} = \left(\frac{m_{\mathrm{pr}}}{m_{\mathrm{P}}}\right)^2 \approx 5.91 \times 10^{-39}, \tag{1.11}$$

where m_{pr} denotes the proton mass. Its smallness is responsible for the unimportance of quantum-gravitational effects on laboratory and astrophysical scales, and for the

[5]'Diese Grössen behalten ihre natürliche Bedeutung so lange bei, als die Gesetze der Gravitation, der Lichtfortpflanzung im Vacuum und die beiden Hauptsätze der Wärmetheorie in Gültigkeit bleiben, sie müssen also, von den verschiedensten Intelligenzen nach den verschiedensten Methoden gemessen, sich immer wieder als die nämlichen ergeben.'

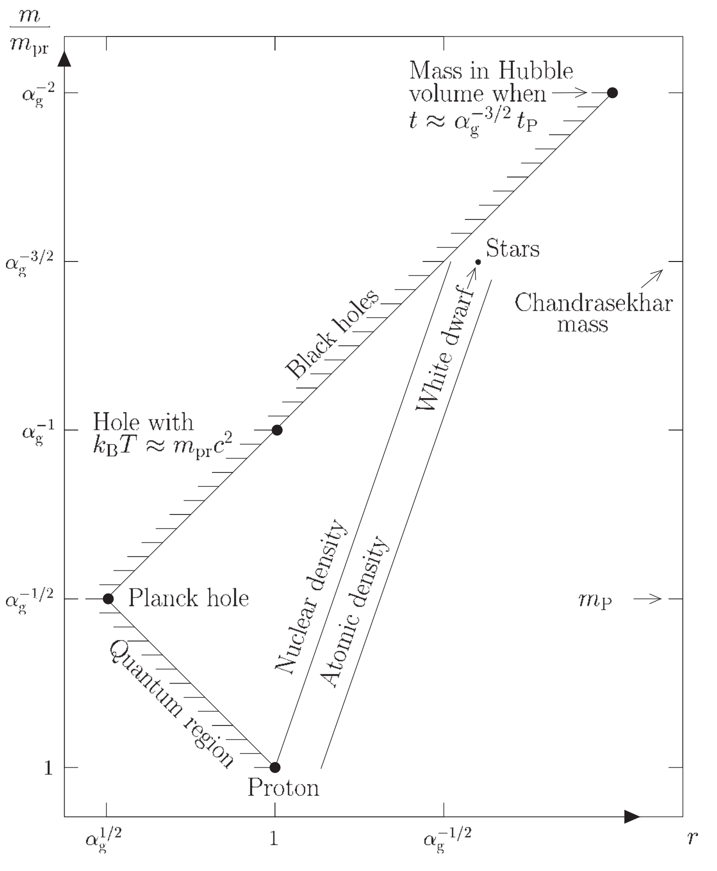

Fig. 1.1 Structures in the Universe (adapted from Rees (1995)).

separation between micro- and macrophysics. As can be seen from the diagram, important features occur for masses that contain simple powers of $\alpha_{\rm g}$ (in terms of $m_{\rm pr}$); cf. Rees (1995). For example, the Chandrasekhar mass $M_{\rm C}$ is given by

$$M_{\rm C} \approx \alpha_{\rm g}^{-3/2} m_{\rm pr} \approx 1.8 M_{\odot}. \tag{1.12}$$

(A more precise value is $M_{\rm C} \approx 1.44 M_{\odot}$. It gives the upper limit for the mass of a white dwarf and sets the scale for stellar masses.) The minimum stellar lifetimes contain $\alpha_{\rm g}^{-3/2} t_{\rm P}$ as the important factor. It is also interesting to note that the geometric mean of the Planck length and the size of the observable part of the Universe is about 0.1 mm—a scale of everyday life. It is an open question whether fundamental theories such as quantum gravity can provide an explanation for such values, for example the ratio $m_{\rm pr}/m_{\rm P}$, or not. Tegmark *et al.* (2006) give a list of 31 dimensionless parameters in particle physics and cosmology that demand a fundamental explanation. We shall come back to this in Chapter 10.

As far as the relationship between quantum theory and the gravitational field is concerned, one can distinguish between different levels. The first, lowest level deals with quantum *mechanics* in *external* gravitational fields (described by either GR or its Newtonian limit). No back reaction on the gravitational field is taken into account. This is the only level where experiments exist so far. The next level concerns quantum *field* theory in *external* gravitational fields described by GR. Back reaction can be taken into account in a perturbative sense. These two levels will be dealt with in the next two subsections. The highest level, *full* quantum gravity, will be discussed in the rest of this book.

1.1.4 Quantum mechanics and Newtonian gravity

Consider first the level of Newtonian gravity. There exist experiments that test the classical trajectories of elementary particles, such as thermal neutrons that fall like mass points; see, for example, Hehl *et al.* (1991). This is not so much of interest here. We are more interested in quantum-mechanical *interference* experiments concerning the motion of neutrons and atoms in external gravitational fields.

Historically, two experiments have been of significance. The experiment by Colella, Overhauser, and Werner ('COW') in 1975 was concerned with neutron interferometry in the gravitational field of the Earth. According to the equivalence principle, an analogous experiment should be possible with neutrons in accelerated frames. Such an experiment was performed by Bonse and Wroblewski in 1983. Details and references can be found in the reviews by Hehl *et al.* (1991) and Werner and Kaiser (1990).

In the following, we shall briefly describe the 'COW' experiment; see Fig. 1.2. A beam of neutrons is split into two parts, such that they can travel at different heights in the terrestrial gravitational field. They are then recombined and sent to detectors. The whole apparatus can be rotated through a varying angle θ around the horizontal axis. The interferences are then measured in dependence on θ. The theoretical description makes use of the Schrödinger equation for neutrons (Hehl *et al.* 1991). The Hamiltonian in the system of the rotating Earth is given by

$$H = \frac{\mathbf{p}^2}{2m_{\rm i}} + m_{\rm g}\mathbf{g}\mathbf{r} - \boldsymbol{\omega}\mathbf{L}. \tag{1.13}$$

We have distinguished here between the inertial mass, $m_{\rm i}$, of the neutron and its (passive) gravitational mass, $m_{\rm g}$, because 'COW' also used this experiment as a test of the equivalence principle. In the last term, $\boldsymbol{\omega}$ and $\mathbf{L}$ denote the angular velocity of

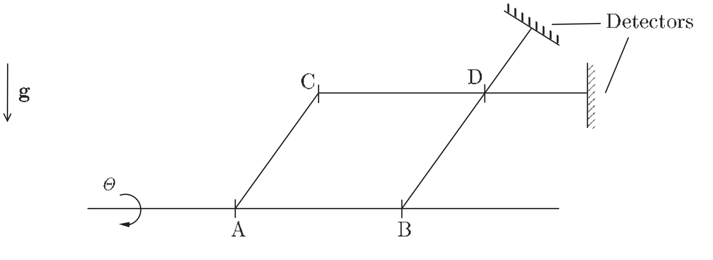

Fig. 1.2 Schematic description of the 'COW' experiment for neutron interferometry in the gravitational field of the Earth.

the Earth and the angular momentum of the neutron with respect to the centre of the Earth (given by $\mathbf{r} = 0$), respectively. This term describes centrifugal and Coriolis forces. Note that the canonical momentum is given by

$$\mathbf{p} = m_{\rm i}\dot{\mathbf{r}} + m_{\rm i}\boldsymbol{\omega} \times \mathbf{r}. \tag{1.14}$$

The phase shift in the interferometer experiment is given by

$$\Delta\beta = \frac{1}{\hbar}\oint \mathbf{p}\,\mathrm{d}\mathbf{r}, \tag{1.15}$$

where the integration runs over the parallelogram ABDC of Fig. 1.2. According to (1.14), there are two contributions to the phase shift. The term containing ω describes the influence of the terrestrial rotation on the interference pattern ('neutron Sagnac effect'). It yields

$$\Delta\beta_{\rm Sagnac} = \frac{m_{\rm i}}{\hbar}\oint (\boldsymbol{\omega} \times \mathbf{r})\mathrm{d}\mathbf{r} = \frac{2m_{\rm i}}{\hbar}\boldsymbol{\omega}\mathbf{A}, \tag{1.16}$$

where $\mathbf{A}$ denotes the normal area vector of the loop ABDC.

Of main interest here is the gravitational part of the phase shift. Since the contributions of the sides $\overline{\rm AC}$ and $\overline{\rm DB}$ cancel, one has

$$\Delta\beta_{\rm g} = \frac{m_{\rm i}}{\hbar}\oint \mathbf{v}\;\mathrm{d}\mathbf{r} \approx \frac{m_{\rm i}(v_0 - v_1)}{\hbar}\overline{\rm AB}, \tag{1.17}$$

where v_0 and v_1 denote the absolute values of the velocities along $\overline{\rm AB}$ and $\overline{\rm CD}$, respectively. From energy conservation one gets

$$v_1 = v_0\sqrt{1 - \frac{2\Delta V}{m_{\rm i}v_0^2}} \approx v_0 - \frac{m_{\rm g}gh_0\sin\,\theta}{m_{\rm i}v_0},$$

where $\Delta V = m_{\rm g}gh_0\sin\theta$ is the potential difference, h_0 denotes the perpendicular distance between $\overline{\rm AB}$ and $\overline{\rm CD}$, and the limit $2\Delta V/m_{\rm i}v_0^2 \ll 1$ (about 10^{-8} in the experiment) has been used. The neutrons are prepared with a de Broglie wavelength $\lambda = 2\pi\hbar/p \approx 2\pi\hbar/m_{\rm i}v_0$ (neglecting the $\boldsymbol{\omega}$ part, since the Sagnac effect contributes only 2% of the effect), attaining a value of about 1.4 Å in the experiment. One then gets for the gravitational phase shift the final result

$$\Delta\beta_{\rm g} \approx \frac{m_{\rm i}m_{\rm g}g\lambda A\sin\theta}{2\pi\hbar^2}, \tag{1.18}$$

where A denotes the area of the parallelogram ABDC. This result has been confirmed by 'COW' with 1% accuracy. The phase shift (1.18) can be rewritten in an alternative form such that only those quantities appear that are directly observable in the experiment (Lämmerzahl 1996). It then reads

$$\Delta\beta_{\rm g} \approx \frac{m_{\rm g}}{m_{\rm i}}\mathbf{g}\mathbf{G}TT', \tag{1.19}$$

where T denotes the flight time of the neutron from A to B, and T' denotes the flight time from A to C or from B to D; $\mathbf{G}$ is the reciprocal lattice vector of the crystal layers (from which the neutrons are scattered in the beam splitter). Now m_g and m_i appear as in the classical theory as a ratio, not as a product. The 'COW' experiment has confirmed the validity of the (weak) equivalence principle in the quantum domain.

Modern tests prefer to use atom interferometry because atoms are easier to handle and the experiments allow tests of higher precision (Lämmerzahl 1996, 1998). In these experiments the flight time T is just a time between laser pulses, which is then the interaction time with the gravitational field; T is chosen by the experimentalist. In one important application, Peters *et al.* (2001) have used atom interferometry to measure the gravitational acceleration g with a resolution of $\Delta g/g \sim 10^{-10}$. By comparing the gravitational accelerations of two rubidium isotopes, as well as of rubidium atoms in different hyperfine ground states, the weak equivalence principle was confirmed with accuracy of 10^{-7} (Fray *et al.* 2004). Proposals have been made that the weak equivalence principle can be tested with atom interferometry up to one part in 10^{17} (Dimopoulos *et al.* 2007).

Experiments with Bose–Einstein condensates in free fall are also of interest (van Zoest *et al.* 2010). With expansion times of up to 1 s for the condensate, one can observe its ultraslow expansion to a macroscopic matter-wave packet, reaching an extension of the order of millimetres. Bose–Einstein condensates in free fall can thus become a promising source of tests of GR using matter-wave interferometry.

Neutrons are also still useful for studying quantum systems in a gravitational field. An experiment with ultracold neutrons has shown that their vertical motion in a gravitational field has discrete energy states, as predicted by the Schrödinger equation (Nesvizhevsky *et al.* 2002). The minimum energy is 1.4×10^{-12} eV, which is much smaller than the ground-state energy of the hydrogen atom. Another experiment has realized a 'neutron whispering gallery', that is, the localization of neutron waves near a curved reflecting surface in analogy to the long-known phenomenon involving sound waves in air (Nesvizhevsky *et al.* 2010).

Since neutrons are fermions, it is from a fundamental point of view more appropriate to address the Dirac equation than the Schrödinger equation. As we shall see, this gives rise to new effects involving spin. In Minkowski space (and cartesian coordinates), the Dirac equation reads

$$\left(\mathrm{i}\gamma^\mu \partial_\mu + \frac{mc}{\hbar}\right)\psi(x) = 0, \tag{1.20}$$

where $\psi(x)$ is a Dirac spinor, and the matrices γ^μ obey the anticommutation relation

$$[\gamma^\mu, \gamma^\nu]_+ := \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}. \tag{1.21}$$

The transformation into an accelerated frame is achieved by replacing partial derivatives with covariant derivatives (see, for example, Gronwald and Hehl (1996)):

$$\partial_\mu \longrightarrow D_\mu := \partial_\mu + \frac{\mathrm{i}}{4}\sigma^{mk}\omega_{\mu mk}, \tag{1.22}$$

where $\sigma^{mk} = \mathrm{i}[\gamma^m, \gamma^k]/2$ is the generator of the Lorentz group, and $\omega_{\mu mk}$ denotes the components of the connection; Latin indices denote anholonomic components, that is, components that are not derivable from a coordinate basis. From the equivalence principle, one would expect that this also gives the appropriate form in curved space–time, where

$$[\gamma^\mu, \gamma^\nu]_+ = 2g^{\mu\nu}. \tag{1.23}$$

For the formulation of the Dirac equation in curved space–time, one has to use the tetrad ('vierbein') formalism, in which a basis $e_n = \{e_0, e_1, e_2, e_3\}$ is chosen at each space–time point. This is the reason why anholonomic components come into play. One can expand the tetrads with respect to the tangent vectors along coordinate lines (the 'holonomic basis') according to

$$e_n = e_n^\mu \partial_\mu. \tag{1.24}$$

Usually one chooses the tetrad to be orthonormal,

$$e_n \cdot e_m := g_{\mu\nu} e_n^\mu e_m^\nu = \eta_{nm} := \mathrm{diag}(-1, 1, 1, 1). \tag{1.25}$$

The reason why one has to go beyond the pure metric formalism is the fact that spinors (describing fermions) are objects whose wave components transform with respect to a two-valued representation of the Lorentz group. One therefore needs a local Lorentz group and local orthogonal frames.

One can define anholonomic Dirac matrices according to

$$\gamma^n := e_\mu^n \gamma^\mu, \tag{1.26}$$

where $e_n^\mu e_\mu^m = \delta_n^m$. This leads to

$$[\gamma^n, \gamma^m]_+ = 2\eta^{nm}. \tag{1.27}$$

The Dirac equation in curved space–time or accelerated frames then reads, using anholonomic components,

$$\left(\mathrm{i}\gamma^n D_n + \frac{mc}{\hbar}\right)\psi(x) = 0. \tag{1.28}$$

In order to study quantum effects of fermions in the gravitational field of the Earth, one specializes this equation to the non-inertial frame of an accelerated and rotating observer, with linear acceleration $\mathbf{a}$ and angular velocity $\boldsymbol{\omega}$ (see e.g. Hehl *et al.* 1991). A non-relativistic approximation with relativistic corrections is then obtained by the standard Foldy–Wouthuysen transformation, decoupling the positive- and negative-energy states. This leads to (writing $\beta \equiv \gamma^0$)

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = H_{\mathrm{FW}}\psi, \tag{1.29}$$

with

$$\begin{aligned} H_{\mathrm{FW}} = {} & \beta mc^2 + \frac{\beta}{2m}\mathbf{p}^2 - \frac{\beta}{8m^3c^2}\mathbf{p}^4 + \beta m(\mathbf{a}\,\mathbf{x}) \\ & -\boldsymbol{\omega}(\mathbf{L}+\mathbf{S}) + \frac{\beta}{2m}\mathbf{p}\frac{\mathbf{a}\,\mathbf{x}}{c^2}\mathbf{p} + \frac{\beta\hbar}{4mc^2}\vec{\Sigma}(\mathbf{a}\times\mathbf{p}) + \mathcal{O}\left(\frac{1}{c^3}\right), \end{aligned} \tag{1.30}$$

where $\vec{\Sigma}$ is three spin matrices defined by $\Sigma^k = \epsilon^{kmn}\sigma_{mn}/2$. In a convenient representation, one has $\vec{\Sigma} = \mathrm{diag}(\vec{\sigma}, \vec{\sigma})$, where $\vec{\sigma}$ are the Pauli matrices.

The interpretation of the various terms in (1.30) is straightforward. The first four terms correspond to the rest mass, the usual non-relativistic kinetic term, the first relativistic correction to the kinetic term, and the 'COW' effect (or its analogue for pure acceleration), respectively. The term $\boldsymbol{\omega}\mathbf{L}$ describes the Sagnac effect, while $\boldsymbol{\omega}\mathbf{S}$ corresponds to a new spin-rotation effect ('Mashhoon effect') that cannot be found from the Schrödinger equation. One can estimate that for typical values of a neutron interferometer experiment, the Mashhoon effect contributes only 10^{-9} of the Sagnac effect. This is very small, but it has been indirectly observed in the nuclear-spin precession of atomic mercury (Mashhoon 1995).

In the presence of torsion, one would get additional terms in (1.30). Therefore, from investigations of atomic spectra, one can get constraints on the torsion. For example, Lämmerzahl (1997) found the following bound for the spatial component K of the axial torsion: $K \leq 1.5 \times 10^{-15}\ \mathrm{m}^{-1}$. As has already been remarked above, torsion may play a significant role in the early Universe. It was estimated that torsion effects should become important at a density (Hehl *et al.* 1976)

$$\rho = \frac{m_{\mathrm{e}}}{\lambda_{\mathrm{e}} l_{\mathrm{P}}^2} \approx 9 \times 10^{48}\ \frac{\mathrm{g}}{\mathrm{cm}^3} \ll \rho_{\mathrm{P}},$$

where m_{e} is the electron mass and λ_{e} its Compton wavelength. Under the assumption of a radiation-dominated universe with the critical density, the density would already be higher than this value for times earlier than about 10^{-22} s, which is much later than the timescale where the inflationary phase is assumed to happen.

We mention that this framework is also of use in the study of a 'generalized' Dirac equation to devise quantum tests of general relativity (Lämmerzahl 1998) and to construct an axiomatic approach to space–time geometry, yielding a Riemann–Cartan geometry (see Audretsch *et al.* 1992). A detailed review of the interaction of mesoscopic quantum systems with gravity is presented in Kiefer and Weber (2005).

If one starts with the Dirac equation (1.20), one gets a different viewpoint on the weak equivalence principle. Since only the one mass m appears in (1.20), it is obvious that different masses m_{i} and m_{g} cannot occur in the non-relativistic expansion; see (1.30). From the relativistic point of view, therefore, it is redundant to introduce different masses into the non-relativistic Hamiltonian (1.13); there should occur only one mass m. The same holds for the non-relativistic expansion of the Klein–Gordon equation, which is briefly discussed at the end of Section 5.4.1.

In concluding this subsection, we want to discuss briefly one important occasion on which GR seemed to play a role in the foundations of quantum mechanics. This is the discussion of the time–energy uncertainty relations by Bohr and Einstein at the sixth Solvay conference, which took place in Brussels in 1930 (cf. Bohr 1949).

Einstein came up with the following counterargument against the validity of this uncertainty relation. Consider a box filled with radiation. A clock controls the opening of a shutter for a short time interval such that a single photon can escape at a fixed time t. The energy E of the photon is, however, also fixed because it can be determined by weighing the box before and after the escape of the photon. It thus seems as if the time–energy uncertainty relation is violated.

In his response to Einstein's attack, Bohr came up with the following arguments. Consider the details of the weighing process, in which a spring is attached to the box;

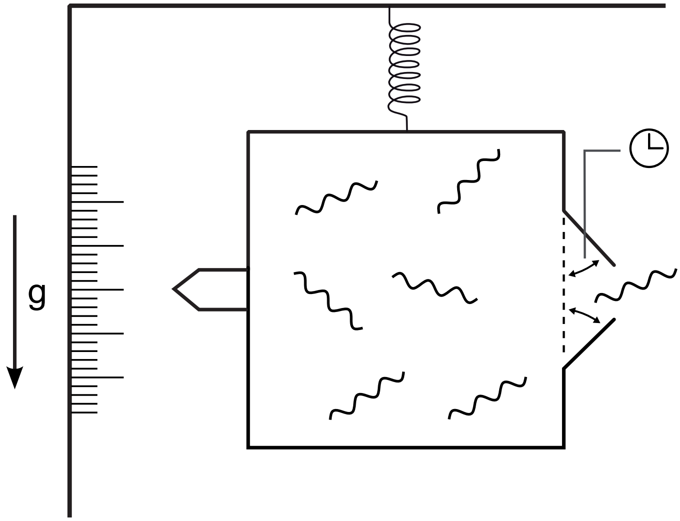

Fig. 1.3 Setting of the gedanken experiment for the Einstein–Bohr debate on the time–energy uncertainty relation.

see Fig. 1.3. The null position of the balance is known with an accuracy Δq. This leads to an uncertainty in the momentum of the box $\Delta p \sim \hbar/\Delta q$. Bohr then makes the assumption that Δp must be smaller than the total momentum imposed by the gravitational field during the time T of the weighing process on the mass uncertainty Δm of the box. This leads to

$$\Delta p < v\Delta m = gT\Delta m, \tag{1.31}$$

where g is the gravitational acceleration. Now GR enters the game: the tick rate of clocks depends on the gravitational potential according to the 'redshift formula'

$$\frac{\Delta T}{T} = \frac{g\Delta q}{c^2}, \tag{1.32}$$

so that, using (1.31), the uncertainty in ΔT after the weighing process is

$$\Delta T = \frac{g\Delta q}{c^2}T > \frac{\hbar}{\Delta mc^2} = \frac{\hbar}{\Delta E}, \tag{1.33}$$

in accordance with the time–energy uncertainty relation. (After this, Einstein gave up trying to find an inconsistency in quantum mechanics, but focused instead on its possible incompleteness.) But are Bohr's arguments really consistent? There are, in fact, some possible loopholes (cf. Shi 2000). First, it is unclear whether (1.31) must really hold, since Δp is an intrinsic property of the apparatus. Second, the relation (1.32) cannot hold in this form, because T is *not* an operator, and therefore ΔT cannot have the same interpretation as Δq. In fact, if T is considered a classical quantity, it would be more consistent to relate Δq to an uncertainty in g, which in fact would suggest considering the quantization of the gravitational field. One can also change the gedanken experiment by using an electrostatic field instead of the gravitational field, where a relation of the form (1.32) no longer holds (see von Borzeszkowski and Treder 1988). It should also be emphasized that not much of GR is needed; in fact, the relation (1.32) follows from energy conservation and $E = h\nu$ alone. A general criticism of all these early gedanken experiments deals with their inconsistent interpretation of measurements as being related to uncontrollable interactions; see Shi (2000) and Chapter 10. The important feature is, however, entanglement between quantum systems.

It thus seems as if Bohr's analysis was mainly based on dimensional arguments. In fact, the usual application of the time–energy uncertainty relation relates line widths of spectra for unstable systems to the corresponding half-life time. In quantum gravity, no time parameter appears on the fundamental level (see Chapter 5). A time–energy uncertainty relation can only be derived in the semiclassical limit.

1.1.5 Quantum field theory in curved space–time

Some interesting new aspects appear when quantum *fields* play a role. They mainly concern the notions of *vacuum* and *particles*. A vacuum is only invariant with respect to Poincaré transformations, so that observers that are not related by inertial motion refer in general to different types of vacua (Fulling 1973). 'Particle creation' can occur in the presence of external fields or for non-inertial observers. An external electric field, for example, can lead to the creation of electron–positron pairs ('Schwinger effect'); see, for example, Grib *et al.* (1994). We shall be mainly concerned with particle creation in the presence of external gravitational fields (Birrell and Davies 1982, Parker and Toms 2009). This was first discussed by Schrödinger (1939).

One example of particular interest is particle creation from black holes (Hawking 1975); see, for example, Frolov and Novikov (1998), Fré *et al.* (1999), and Hehl *et al.* (1998) for a detailed review. This is not only of fundamental theoretical interest, but could also lead to observational consequences. A black hole radiates with a universal temperature ('Hawking temperature') according to

$$T_{\rm BH} = \frac{\hbar\kappa}{2\pi k_{\rm B}c}, \tag{1.34}$$

where κ is the surface gravity of a stationary black hole, which by the no-hair theorem is uniquely characterized by its mass M, its angular momentum J, and (if present) its electric charge q. In the particular case of the spherically symmetric Schwarzschild black hole, one has $\kappa = c^4/4GM = GM/R_{\rm S}^2$ and therefore

$$T_{\rm BH} = \frac{\hbar c^3}{8\pi k_{\rm B}GM} \approx 6.17\times 10^{-8}\left(\frac{M_\odot}{M}\right)\ {\rm K}. \tag{1.35}$$

This temperature is unobservably small for solar-mass (and bigger) black holes, but may be observable for primordial black holes. It must be emphasized that the expression for $T_{\rm BH}$ contains all fundamental constants of nature. One may speculate that this expression—relating the macroscopic parameters of a black hole to thermodynamic quantities—will play a similar role for quantum gravity to that which de Broglie's relations $E = \hbar\omega$ and $p = \hbar k$ once played for the development of quantum theory (Zeh 2007).

Hawking radiation was derived in the semiclassical limit in which the gravitational field can be treated classically. According to (1.35), the black hole loses mass through its radiation and becomes hotter. After it has reached a mass of the size of the Planck mass (1.6), the semiclassical approximation breaks down and the full theory of quantum gravity should be needed. Black-hole evaporation thus plays a crucial role in any approach to quantum gravity; cf. Chapter 7.

There exists an effect related to (1.34) in flat Minkowski space. An observer with uniform acceleration a experiences the standard Minkowski vacuum not as empty, but as filled with *thermal* radiation with temperature

$$T_{\mathrm{DU}} = \frac{\hbar a}{2\pi k_{\mathrm{B}} c} \approx 4.05 \times 10^{-23}\ a \left[\frac{\mathrm{cm}}{\mathrm{s}^2}\right]\ \mathrm{K}. \tag{1.36}$$

This temperature is often called the 'Davies–Unruh temperature' after the work by Davies (1975) and Unruh (1976), with important contributions also by Fulling (1973). Formally, it arises from (1.34) through the substitution of κ by a. This can be understood from the fact that *horizons* are present in both the black-hole case and the acceleration case; see, for example, Kiefer (1999) for a detailed review. Although (1.36) seems to be a small effect, people have suggested looking for it in accelerators (Leinaas 2002) or in experiments with ultraintense lasers where high accelerations can be attained (Thirolf *et al.* 2009).

A central role in the theory of quantum fields on an external space–time is played by the semiclassical Einstein equations. These equations are obtained by replacing the energy–momentum tensor in (1.1) by the expectation value of the energy–momentum operator with respect to some quantum state Ψ,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} \langle \Psi | \hat{T}_{\mu\nu} | \Psi \rangle. \tag{1.37}$$

A particular issue is the regularization and renormalization of the object on the right-hand side (Birrell and Davies 1982). This leads, for example, to a flux of negative energy into the black hole, which can be interpreted as the origin of Hawking radiation. As we shall discuss in the next section, (1.37) is of limited value if seen from the viewpoint of the full quantum theory. We shall find in Section 5.4 that (1.37) can be derived approximately from canonical quantum gravity as a kind of mean-field equation.

1.2 Problems of a fundamentally semiclassical theory

1.2.1 Does one really have to quantize gravity?

When one is dealing with approaches to quantum gravity, the question is sometimes asked whether it is really necessary to quantize the gravitational field. And even if it is, doubts have occasionally been expressed whether such a theory can operationally be distinguished from an 'exact' semiclassical theory.[6] As a candidate for the latter, the semiclassical Einstein equations (1.37) are often presented; cf. Møller (1962). The

[6] 'Semiclassical' here means an exact theory that couples quantum degrees of freedom to classical degrees of freedom. It therefore has nothing to do with the WKB approximation, which is usually referred to as the semiclassical approximation. The latter is discussed in Section 5.4.

more general question behind this issue concerns the possibility of a consistent *hybrid dynamics* through which a quantum and a classical system can be coupled.

Can one present general arguments that would *enforce* the quantization of the gravitational field? Bryce DeWitt wrote (DeWitt 1962, p. 272)

> It is shown in a quite general manner that the quantization of a given system implies also the quantization of any other system to which it can be coupled.

Moreover, Eppley and Hannah (1977) argued that the coupling of a classical gravitational wave with a quantum system would lead to inconsistencies. As shown in detail in Albers *et al.* (2008), it is not possible to draw such strong conclusions. Let us repeat here the main arguments.

Eppley and Hannah consider the interaction of a classical gravitational wave of small momentum with a quantum system of one particle described by a wave function ψ. They restrict their considerations to the case in which the wavelength of the wave is much smaller than the position uncertainty of the particle in order for the interaction to lead to a measurement of the position of the particle. They then distinguish between two possibilities: either the gravitational wave leads to a collapse of the wave function of the particle, or it does not. If it does, they argue that this would entail either momentum non-conservation or a violation of the uncertainty relation. If it does not collapse, they argue that there will be a superluminal transmission of information.

Let us first consider the case of an assumed collapse. Such a collapse is not part of standard quantum theory, because it would be in conflict with the superposition principle. There exist various models of an explicit wave-function collapse in the literature; see, for example, Chapter 8 in Joos *et al.* (2003). These models often have problems with conservation laws, so the occurrence of momentum non-conservation in the gedanken experiment of Eppley and Hannah is not surprising and should be taken as an argument against the collapse situation, but not as an argument in favour of quantizing gravity. Moreover, there is so far no experimental support for any collapse model.

Let us then assume that the gravitational wave does not collapse the wave function of the particle. Eppley and Hannah consider here an Einstein–Podolsky–Rosen (EPR) type situation where one particle decays into two other particles (say, two photons) which together are in a singlet state. These authors then argue that a classical gravitational wave scattered off photon 2 can distinguish between photon 2 having a definite polarization or photon 2 being in a superposition of both polarizations. If a measurement is done on photon 1, so the argument goes, the information about it will instantaneously arrive at the gravitational wave, corresponding to superluminal speed.

This conclusion is, however, not correct. Before the measurement of photon 1, the total state of the photons and the detector is the entangled state

$$|\Psi_0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2\right)|\Phi_0\rangle, \tag{1.38}$$

where $|\uparrow\rangle_1$ and $|\downarrow\rangle_1$ denote horizontal and vertical polarizations, respectively, of photon 1 (and similarly for photon 2), and $|\Phi_0\rangle$ denotes the initial (switched-off) state of the detector which will measure photon 1. After the detector has measured photon 1, the initial state $|\Psi_0\rangle$ will have evolved into the new entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_1 |\downarrow\rangle_2 |\Phi_\uparrow\rangle - |\downarrow\rangle_1 |\uparrow\rangle_2 |\Phi_\downarrow\rangle \right), \tag{1.39}$$

where $|\Phi_\uparrow\rangle$ and $|\Phi_\downarrow\rangle$ denote the state of the detector after it has measured the polarization of photon 1 to be horizontal or vertical, respectively. The important point is that photon 2, by itself, is not in a pure state, neither before nor after the measurement of photon 1. It finds itself in a mixed state, which is obtained from the total pure state of system plus detector by tracing out the states of particle 1 and the detector. This leads for *both* (1.38) and (1.39) to the same density operator for photon 2,

$$\hat{\rho} = \frac{1}{2} \left(|\uparrow\rangle_2 \langle\uparrow|_2 + |\downarrow\rangle_2 \langle\downarrow|_2 \right). \tag{1.40}$$

In both cases, photon 2, by itself, is in a mixed state of horizontal and vertical polarization with equal probabilities of one half. The gravitational wave thus 'sees' the same mixed state for photon 2, independent of whether a measurement of photon 1 has been performed or not; thus, no superluminal communication is possible. The difference between the total states (1.38) and (1.39) can only be seen after the two photons are brought together: in case (1.38) they will interfere, while in case (1.39) they will not; in the latter case the information about the original superposition has been delocalized into an entangled quantum state of the detector state with its natural environment; that is, decoherence has occurred (see Chapter 10).

This chain of arguments leaves the linear structure of quantum theory untouched, which corresponds to an Everett interpretation (see Chapter 10). If we assume instead that a measurement on photon 1 collapses its wave function into $|\uparrow\rangle_1$ or $|\downarrow\rangle_1$, a superluminal communication might in principle be possible (Carlip 2008). This can happen, for example, in the context of the semiclassical Einstein equations (1.37), where the source of the gravitational field is taken to be the quantum expectation value of the energy–momentum tensor. Semiclassical gravity introduces a non-linearity into quantum theory, which in principle can be experimentally tested (Carlip 2008). The possibility of such an example does not prove, however, that a mixed classical–quantum coupling without superluminal communication is impossible. One example is presented in Albers *et al.* (2008). It is based on the general framework developed by Hall and Reginatto (2005) using ensembles in configuration space and contains a consistent coupling between classical gravity (described for simplicity by a scalar field) and a matter scalar field. This example demonstrates that it is possible to devise a classical–quantum coupling without superluminal communication.

What, then, about DeWitt's quote above that the quantization of a given system implies also the quantization of any other system to which it is coupled? As shown in Albers *et al.* (2008), it is certainly true that the uncertainties present in one system entail uncertainties in another system to which it is coupled. But these uncertainties can be entirely classical, and the formalism developed by DeWitt is, in fact, general enough to encompass both classical and quantum systems. It cannot be concluded from this analysis alone that the validity of quantum uncertainty relations in one system entails the validity of such relations in the coupled system. Again, the example of the quantum–classical system presented in Albers *et al.* (2008) is an explicit counterexample to DeWitt's claim.

The above arguments show that no logical argument can force the quantization of a system that is coupled to a quantum system. If we believe in the universality of quantum theory, however, the quantization of gravity will be unavoidable. Feynman, in a discussion contribution to the Chapel Hill Conference in 1957, expressed this point very clearly; see DeWitt (1957), pp. 136–42, and also Zeh (2011). He considered a Stern–Gerlach experiment in which an electron is deflected as usual to an upper and a lower counter, depending on the value of the corresponding spin component. In addition, the counters are supposed to be connected to a little ball 1 cm in diameter, which is deflected upwards or downwards depending on the spin value. If the electron is in a superposition of spin up and spin down, the ball will also be in a superposition of up and down—provided quantum theory is valid up to that level. But then the gravitational field of the ball will also be in a superposition; that is, gravity will be quantized. As Feynman says:

> ...if you believe in quantum mechanics up to any level then you have to believe in gravitational quantization in order to describe this experiment. ...It may turn out, since we've never done an experiment at this level, that it's not possible ...that there is something the matter with our quantum mechanics when we have too much *action* in the system, or too much mass—or something. But that is the only way I can see which would keep you from the necessity of quantizing the gravitational field. It's a way that I don't want to propose. ...

1.2.2 A Bohr–Rosenfeld type of analysis

It is often argued that the famous gedanken experiments by Bohr and Rosenfeld (1933) imply that a coupling between a classical and a quantum system is inconsistent. For this reason, we shall briefly review their arguments here (see also Heitler (1984) for a lucid discussion).

Historically, Landau and Peierls (1931) had claimed that the quantum nature of the *electromagnetic field* cannot be tested, since in their investigation they found that there exists a fundamental minimal uncertainty for single field amplitudes, not only for conjugate pairs. Bohr and Rosenfeld then showed that this is not true. Their line of thought runs as follows. Consider a charged body with mass M and charge Q, acting as a measuring device for the electric field that is present in a volume $V \equiv l^3$. Momentum measurements of the body are made at the beginning and the end of the time interval of the measurement. In order to qualify the body as a measurement device, the following assumptions are made (in order to avoid back reaction, etc.) (see also von Borzeszkowski and Treder 1988):

$$Q^2 \gtrsim \hbar c \approx 137 e^2, \tag{1.41}$$

$$l > \frac{Q^2}{Mc^2}. \tag{1.42}$$

The latter condition expresses the fact that the electrostatic energy should be smaller than the rest mass. Bohr and Rosenfeld then found from their detailed analysis the following necessary conditions, with $\mathcal{E}$ denoting the average of an electric field component over the volume V:

$$\Delta\mathcal{E}\, l^2 \gtrsim \frac{\hbar c}{Q}, \tag{1.43}$$

$$\Delta\mathcal{E}\, l^3 \gtrsim \frac{\hbar Q}{Mc}. \tag{1.44}$$

One can now always choose a measurement device such that the ratio Q/M in the last expression can be made arbitrarily small. Therefore, $\mathcal{E}$ can be measured with arbitrary precision, contrary to the arguments of Landau and Peierls (1931). Bohr and Rosenfeld then showed that (1.43) and (1.44) are in agreement with the uncertainty relations derived from the quantum commutators of quantum electrodynamics (QED). The inaccuracy of the field produced by the test body is responsible therein for the limitations in the simultaneous measurement of conjugate field quantities. Their discussion, therefore, shows the *consistency* of the formalism with the measurement analysis. It does *not* provide a logical proof that the electromagnetic field must be quantized; cf. Rosenfeld (1963).[7] Or, as Amelino-Camelia and Stachel (2009) have put it, '...the limits of *definability* of a quantity within any formalism should coincide with the limits of *measurability* of that quantity for all conceivable (ideal) measurement procedures.'

Although the final formalism for quantum gravity is not at hand, the Bohr–Rosenfeld analysis can at least formally be extended to the gravitational field; cf. Bronstein (1936), DeWitt (1962), and von Borzeszkowski and Treder (1988). One can replace the electric field $\mathcal{E}$ with the Christoffel symbols Γ (the 'gravitational force'). Since one can then perform in the Newtonian approximation, where $\Gamma \sim GM_{\rm g}/r^2c^2$ ($M_{\rm g}$ denoting the gravitational mass), the substitutions

$$\Delta\mathcal{E} \to \frac{c^2\Delta\Gamma}{G}, \quad Q \to M_{\rm g}, \tag{1.45}$$

one gets from (1.44) the relation (writing $M_{\rm i}$ instead of M to emphasize that it is the inertial mass)

$$\Delta\Gamma\, l^3 \gtrsim \frac{\hbar G}{c^3}\frac{M_{\rm g}}{M_{\rm i}}. \tag{1.46}$$

Using the (weak) equivalence principle, $M_{\rm g} = M_{\rm i}$, and recalling the definition (1.4) of the Planck length, one can write

$$\Delta\Gamma \gtrsim \frac{l_{\rm P}^2}{l^3}. \tag{1.47}$$

The analogous relation for the metric g would then read

$$\Delta g \gtrsim \left(\frac{l_{\rm P}}{l}\right)^2. \tag{1.48}$$

Thus, the measurement of a single quantity (the metric) seems to be operationally restricted.[8] This is, of course, possible because, unlike the case in QED, the fundamental length scale $l_{\rm P}$ is available. On the other hand, a gedanken experiment by

[7] From *empirical* arguments, we know, of course, that the electromagnetic field is of quantum nature.

[8] Equation (1.48) is similar to the heuristic relation $\Delta g \gtrsim l_{\rm P}/l$ of Misner *et al.* (1973), although the exponent is different.

Smith and Bergmann (1979) shows that the magnetic-type components of the Weyl tensor in linearized quantum gravity can be measured, provided a suitable average over space–time domains is performed.

In a variant of the Bohr–Rosenfeld analysis, Bronstein (1936) found the following limitation on the measurability of the Christoffel symbols:

$$\Delta\Gamma \gtrsim \frac{1}{c^2 T}\left(\frac{\hbar^2 G}{c\rho V^2}\right)^{1/3},$$

where ρ is the mass density of the test body, and an averaging is performed over a time period T. In the limit $\rho \to \infty$, no restriction on $\Delta\Gamma$ would result. However, as has been noticed by Bronstein, the dimensions of the body cannot be smaller than its Schwarzschild radius, that is,

$$\frac{G\rho V}{c^2} < V^{1/3}.$$

Writing again $V \sim l^3$, one then gets again a limitation on the measurability of Γ,

$$\Delta\Gamma \gtrsim \frac{1}{c^2 T}\left(\frac{\hbar^2 G^2}{c^3 l^4}\right)^{1/3} = \frac{1}{cT}\left(\frac{l_{\rm P}}{l}\right)^{4/3}.$$

Bronstein concluded from this limitation that there are fundamental limits on the notion of a Riemannian geometry. He writes in another paper (the quotation is from Gorelik (2005, p. 1048)),

> Matters are different in the quantum theory of the gravitational field. It has to be taken into consideration the limitation arising because the gravitational radius of the testbody ...cannot be larger than its real linear dimension ...measurements of the gravitational field values may be regarded as 'predictable' only if the consideration is restricted to sufficiently large volumes and time intervals. The elimination of the logical inconsistencies connected with this requires a radical reconstruction of the theory, and in particular, the rejection of a Riemannian geometry dealing, as we see here, with values unobservable in principle, and perhaps also the rejection of our ordinary concepts of space and time, modifying them by some much deeper and nonevident concepts. *Wer's nicht glaubt, bezahlt einen Taler.*

A similar viewpoint was taken by Solomon (1938). For a historical account of early work on measurement analysis, see Gorelik (2005) and Stachel (1999); for an extended discussion of minimal lengths obtained from quantum gravity, see Garay (1995).

Does (1.48) imply that the quantum nature of the gravitational field cannot be tested? Not necessarily, for the following reasons. First, there might exist other measurement devices which do not necessarily obey the above relations. Second, this analysis does not say anything about global situations (black holes, cosmology) and about non-trivial applications of the superposition principle. And third, various approaches to quantum gravity seem to predict the existence of a smallest scale; see Chapter 6 on loop quantum gravity and Chapter 9 on string theory; see also Garay (1995). Then, relations such as (1.48) could be interpreted as a confirmation of quantum gravity. It might, of course, be possible, as argued in von Borzeszkowski and Treder (1988) (see also Dyson 2004), that quantum-gravitational analogues of effects such as Compton scattering or the Lamb shift are unobservable in the laboratory, and that only astrophysical tests could be feasible. But such tests could nevertheless reveal the quantum nature of gravity.

1.2.3 The semiclassical Einstein equations

Returning to the specific equations (1.37) for a semiclassical theory, that is, the semiclassical Einstein equations, there are a number of problems attached to them. First, the expectation value of the energy–momentum tensor that occurs on the right-hand side is usually divergent and needs some regularization and renormalization. Such a procedure leads to an essentially unique result for $\langle \hat{T}_{\mu\nu} \rangle$ if certain physical requirements are imposed; cf. Birrell and Davies (1982). The ambiguities can then be absorbed by a redefinition of constants appearing in the action; cf. Section 2.2.3. In this process, however, counterterms arise that invoke higher powers of the curvature such as R^2, which may alter the semiclassical equations at a fundamental level. A possible consequence could be the instability of Minkowski space; see, for example, Ford (2005) and the references therein. In certain exotic situations, $\langle \hat{T}_{\mu\nu} \rangle$ may not even exist, for example on a chronology horizon (Visser 2003).

Second, (1.37) introduces the following element of non-linearity. The space–time metric g depends on the quantum state in a complicated way, since in (1.37) $|\Psi\rangle$ also depends on g through the (functional) Schrödinger equation (an equivalent statement holds in the Heisenberg picture). Consequently, if g_1 and g_2 correspond to states $|\Psi_1\rangle$ and $|\Psi_2\rangle$, respectively, there is no obvious relation between the metric corresponding to a superposition $A|\Psi_1\rangle + B|\Psi_2\rangle$ (which still satisfies the Schrödinger equation) and the metrics g_1 and g_2. This was already remarked on by Anderson (Møller 1962) and by Belinfante in a discussion with Rosenfeld (see Infeld 1964). According to von Borzeszkowski and Treder (1988), it was also the reason why Dirac strongly objected to (1.37).

Rosenfeld insisted on (1.37) because he strongly followed Bohr's interpretation of the measurement process, for which classical concepts should be indispensable (the 'Copenhagen interpretation'). This holds in particular for the structure of space–time, so he wished to have a c-number representation for the metric. He rejected a quantum description for the total system and answered to Belinfante (Infeld 1964) that Einstein's equations may merely be thermodynamical equations of state that break down for large fluctuations, that is, the gravitational field may only be an effective, not a fundamental, field; cf. also Jacobson (1995).

The problem with the superposition principle can be demonstrated by the following argument that has even been put to an experimental test (Page and Geilker 1981).[9] One assumes that there is no explicit collapse of $|\Psi\rangle$, because otherwise one would expect the covariant conservation law $\langle \hat{T}_{\mu\nu} \rangle ;^{\nu} = 0$ to be violated, in contradiction to (1.37). If the gravitational field were quantized, one would expect that each component of the superposition in $|\Psi\rangle$ would act as a source for the gravitational field. This is of course the Everett interpretation of quantum theory; cf. Chapter 10. On the other hand, the semiclassical Einstein equations (1.37) depend on *all* components of $|\Psi\rangle$ simultaneously. Page and Geilker (1981) envisaged the following gedanken experiment, reminiscent of Schrödinger's cat, to distinguish between these options; see Fig. 1.4.

[9] Because in the Copenhagen interpretation the wave function stands merely for the amount of human knowledge about a quantum system, Rosenfeld would have most likely not followed the arguments of Page and Geilker.

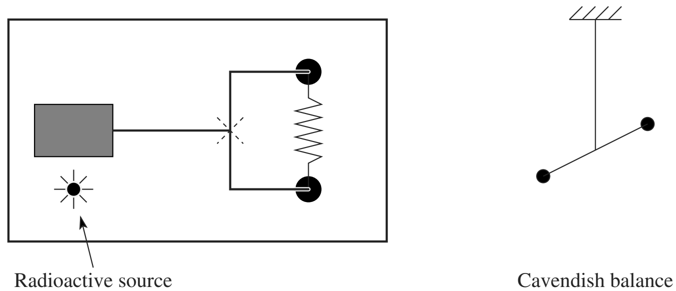

Fig. 1.4 The Page–Geilker experiment.

In a box, there is a radioactive source together with two masses that are connected by a spring. Initially, the masses are rigidly connected, so that they cannot move. If a radioactive decay happens, the rigid connection will be broken and the masses can swing towards each other. Outside the box, there is a Cavendish balance that is sensitive to the location of the masses and therefore acts as a device to 'measure' their position. Following Unruh (1984), the situation can be described by the following simple model. We denote by $|0\rangle$ the quantum state of the masses with a rigid connection, and by $|1\rangle$ the corresponding state in which they can move towards each other. For the purpose of this experiment, it is sufficient to go to the Newtonian approximation of GR and to use the Hamilton operator $\hat{H}$ instead of the full energy–momentum tensor $\hat{T}_{\mu\nu}$. At the initial time $t = 0$, it is assumed that the state is given by $|0\rangle$. For $t > 0$, the state then evolves into a superposition of $|0\rangle$ and $|1\rangle$,

$$|\Psi\rangle(t) = \alpha(t)|0\rangle + \beta(t)|1\rangle,$$

with the coefficients $|\alpha(t)|^2 \approx \mathrm{e}^{-\lambda t}$, $|\beta|^2 \approx 1 - \mathrm{e}^{-\lambda t}$, according to the law of radioactive decay, with a decay constant λ. From this, one finds for the evolution of the expectation value

$$\langle\Psi|\hat{H}|\Psi\rangle(t) = |\alpha(t)|^2\langle 0|\hat{H}|0\rangle + |\beta(t)|^2\langle 1|\hat{H}|1\rangle + 2\,\mathrm{Re}\left[\alpha^*\beta\langle 0|\hat{H}|1\rangle\right].$$

If one makes the realistic assumption that the states are approximate eigenstates of the Hamiltonian, the last term, which describes interferences, vanishes. Anyway, this is not devised as an interference experiment (in contrast to Schrödinger's cat), and interferences would become small due to decoherence (Chapter 10). One is thus left with

$$\langle\Psi|\hat{H}|\Psi\rangle(t) \approx \mathrm{e}^{-\lambda t}\langle 0|\hat{H}|0\rangle + \left(1 - \mathrm{e}^{-\lambda t}\right)\langle 1|\hat{H}|1\rangle. \tag{1.49}$$

According to semiclassical gravity as described by (1.37), therefore, the Cavendish balance would follow the dynamics of the expectation value and swing slightly in the course of time. This is in sharp contrast to the prediction of linear quantum gravity, where in each component the balance reacts to the mass configuration and would thus be observed to swing instantaneously at a certain time. This is, in fact, what has been observed in an actual experiment (Page and Geilker 1981). This experiment, albeit simple, demonstrates that (1.37) cannot fundamentally be true under the given assumption (validity of the Everett interpretation). If one is unbiased with regard to the invoked interpretation, one can continue to search for experimental tests of (1.37). Since these equations are very complicated, one can try to address the 'Schrödinger–Newton equation'

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi - m\Phi\psi \ , \quad \nabla^2\Phi = 4\pi Gm|\psi|^2,$$

which leads to a non-linear and non-local self-interaction in the Schrödinger equation,

$$\mathrm{i}\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{x},t) - Gm^2\int \mathrm{d}^3y\ \frac{|\psi(\mathbf{y},t)|^2}{|\mathbf{x}-\mathbf{y}|}\psi(\mathbf{x},t).$$

This is a restricted version of (1.37) in the Newtonian limit. The additional self-interaction can lead to a decrease in the dispersion of a wave packet, with the amount of decrease becoming stronger with increasing mass m. Numerical estimates indicate, however, that an experimental test of the Schrödinger–Newton equation is not possible in the near future (Giulini and Großardt 2011).

A fundamental non-linear equation such as (1.37) could have far-reaching consequences. It has been remarked, for example, that the validity of the semiclassical Einstein equations (or some non-linear theory of quantum gravity) could be used to break the perfect security of protocols in quantum cryptography (Plaga 2006).

In the Page–Geilker experiment, the reason for the deviation between the predictions of the semiclassical theory and the 'full' theory lies in the large fluctuation for the Hamiltonian. In fact, the experiment was devised to generate such a case. Large fluctuations also occur in another interesting situation—the gravitational radiation emitted by quantum systems (Ford 1982). The calculations are performed for linearized gravity, that is, for a small metric pertubation around flat space–time with metric $\eta_{\mu\nu}$; see, for example, Misner *et al.* (1973) and Chapter 2. Denoting by $G_r(x,x')$ the retarded Green function, one finds for the *integrated* energy–momentum tensor $S_{\mu\nu}$ in the semiclassical theory described by (1.37), the expression[10]

$$\begin{aligned} S^{\mu\nu}_{\mathrm{sc}} = -\frac{8\pi G}{c^4}\int & \mathrm{d}^3x\,\mathrm{d}^4x'\,\mathrm{d}^4x''\ \partial^\mu G_r(x,x')\partial^\nu G_r(x,x'') \\ & \times \left[\langle T_{\alpha\beta}(x')\rangle\langle T^{\alpha\beta}(x'')\rangle - \tfrac{1}{2}\ \langle T(x')\rangle\langle T(x'')\rangle\right] , \end{aligned} \tag{1.50}$$

where $T := T^{\mu\nu}\eta_{\mu\nu}$ denotes the trace of the energy–momentum tensor. On the other hand, quantization of the linear theory (see Chapter 2) yields

$$\begin{aligned} S^{\mu\nu}_{\mathrm{q}} = -\frac{8\pi G}{c^4}\int & \mathrm{d}^3x\,\mathrm{d}^4x'\,\mathrm{d}^4x''\ \partial^\mu G_r(x,x')\partial^\nu G_r(x,x'') \\ & \times \left\langle T_{\alpha\beta}(x')T^{\alpha\beta}(x'') - \tfrac{1}{2}\ T(x')T(x'')\right\rangle . \end{aligned} \tag{1.51}$$

The difference in these results can easily be interpreted: in the semiclassical theory, $\langle T_{\mu\nu}\rangle$ acts as a source, and so no two-point functions $\langle T\ldots T\rangle$ can appear, in contrast to linear quantum theory.

[10] Hats on operators are avoided for simplicity.

It is obvious that the above two expressions differ strongly once the fluctuation of the energy–momentum tensor is large. As a concrete example, Ford (1982) takes a massless real scalar field as a matter source. For coherent states, there is no difference between (1.50) and (1.51). This is not unexpected, since coherent states are as 'classical' as possible, and so the semiclassical and the full theory give identical results. For a superposition of coherent states, however, this is no longer true, and the energies emitted by the quantum system via gravitational waves can differ by macroscopic amounts. For example, if the scalar field is in an eigenstate of the number operator, the semiclassical theory does not predict any radiation at all ($\langle T_{\mu\nu} \rangle$ is time-independent), whereas there is radiation in quantum gravity ($\langle T_{\mu\nu} T_{\rho\lambda} \rangle$ is time-dependent).[11] Therefore, one can in principle have macroscopic quantum-gravity effects even far away from the Planck scale! This is a direct consequence of the superposition principle, which in a linear theory of quantum gravity is valid independent of a particular scale.

Kuo and Ford (1993) have extended this analysis to situations where the expectation value of the energy density can be negative. They show that in such cases the fluctuations in the energy–momentum tensor are large and that the semiclassical theory gives different predictions than the quantum theory. This is true, in particular, for a squeezed vacuum state describing particle creation—a case that is relevant, for example, to structure formation in the Universe; see the remarks in Section 10.1.3. Another example is the Casimir effect. Kuo and Ford (1993) show that the gravitational field produced by the Casimir energy is *not* described by a fixed classical metric.

It will be discussed in Section 5.4 to what extent the semiclassical equations (1.37) can be derived as approximations from full quantum gravity. Modern developments in quantum mechanics discuss the possibility of a consistent formulation of 'hybrid dynamics', coupling a quantum to a classical system; see, for example, Diósi *et al.* (2000). This leads to equations that generalize mean-field equations such as (1.37). It seems that such a coupling can be formulated consistently if the 'classical' system is, in fact, a decohered quantum system (Halliwell 1998). However, this already refers to an effective and not to a fundamental level of description.

One consistent model coupling a classical gravitational field with quantum matter is the one mentioned at the end of Subsection 1.2.1. It is based on a canonical formalism for describing statistical ensembles on configuration space (Hall and Reginatto 2005, Albers *et al.* 2008); whether such a model could play a role in a fundamental theory is, however, far from clear.

1.3 Approaches to quantum gravity

As we have seen in the preceding sections, there exist strong arguments supporting the idea that the gravitational field is of *quantum* nature at the fundamental level. The major task, then, is the construction of a consistent quantum theory of gravity that can be subject to experimental tests.

Can one get hints how to construct such a theory from observation? The idea of a direct probe of the Planck scale (1.6) in high-energy experiments is illusory. In fact, an accelerator of current technology would have to be of the size of several thousand light

[11] Analogous results hold for electrodynamics, with the current j_μ instead of $T_{\mu\nu}$.

years in order to probe the Planck energy $m_{\rm P}c^2 \approx 10^{19}$ GeV. However, we have seen in Section 1.2 that macroscopic effects of quantum gravity could in principle occur at lower energy scales, and we will encounter some other examples in the course of this book. Among these, there are effects of the full theory such as non-trivial applications of the superposition principle for the quantized gravitational field or the existence of discrete quantum states in black-hole physics or the early universe. But one might also be able to observe quantum-gravitational correction terms to established theories, such as correction terms to the functional Schrödinger equation in an external space–time, or effective terms violating the weak equivalence principle. Such effects could potentially be measured in the anisotropy spectrum of the cosmic microwave background radiation or in the forthcoming satellite tests of the equivalence principle such as the missions MICROSCOPE and STEP.

One should also keep in mind that the final theory (which is not yet available) will make its own predictions, some perhaps in a totally unexpected direction. As Heisenberg recalls from a conversation with Einstein[12]:

> From a fundamental point of view it is totally wrong to aim at basing a theory only on observable quantities. For in reality it is just the other way around. Only the theory decides about what can be observed.

A really fundamental theory should have such a rigid structure that all phenomena in the low-energy regime, such as particle masses or coupling constants, can be predicted in a unique way. As there is no direct experimental hint yet, most work in quantum gravity focuses on the attempt to construct a mathematically and conceptually consistent (and appealing) framework.

There is, of course, no a priori given starting point in the methodological sense. In this context, Isham (1987) makes a distinction between a 'primary theory of quantum gravity' and a 'secondary theory'. In the primary approach, one starts with a given classical theory and applies heuristic quantization rules. This is the approach usually adopted, which was successful, for example, in QED. Often the starting point is general relativity, leading to 'quantum general relativity' or 'quantum geometrodynamics', but one could also start from another classical theory such as the Brans–Dicke theory. One usually distinguishes between canonical and covariant approaches. The former employs at the classical level a split of space–time into space and time, whereas the latter aims at preserving four-dimensional covariance at each step. They will be discussed in Chapters 5 and 6, and in Chapter 2, respectively. The main advantage of these approaches is that the starting point is given. The main disadvantage is that one does not arrive immediately at a unified theory of all interactions.

The opposite holds for a 'secondary theory'. One would like to start with a fundamental quantum framework of all interactions and try to derive (quantum) general relativity in certain limiting situations, for example, through an energy expansion. In such a secondary theory, quantum gravity would be an emergent phenomenon. The most important example here is string theory (see Chapter 9). The main advantage is that the fundamental quantum theory automatically yields a unification, a 'theory

[12]'Aber vom prinzipiellen Standpunkt aus ist es ganz falsch, eine Theorie nur auf beobachtbare Größen gründen zu wollen. Denn es ist ja in Wirklichkeit genau umgekehrt. Erst die Theorie entscheidet darüber, was man beobachten kann.' (Einstein according to Heisenberg (1985, p. 92))

of everything'; cf. Weinberg (1993). The main disadvantage is that the starting point is entirely speculative. Short reviews of the main approaches to quantum gravity have been given, for example, by Carlip (2001), Esposito (2011), and Kiefer (2006).

In this book, we shall mainly focus on quantum GR because it is closer to established theories and because it exhibits many general aspects more clearly. In any case, even if quantum GR is superseded by a more fundamental theory such as string theory (which is not obvious), it should be valid as an *effective theory* in some appropriate limit. The reason is that far away from the Planck scale, classical general relativity is the appropriate theory, which in turn must be the classical limit of an underlying quantum theory. Except perhaps close to the Planck scale itself, quantum GR should be a viable framework (like QED, which is also supposed to be only an effective theory). It should also be emphasized that string theory automatically implements many of the methods used in the primary approach, such as quantization of constrained systems and covariant perturbation theory.

An important question in the heuristic quantization of a given classical theory is which of the classical structures should be subjected to the superposition principle and which should remain classical (or absolute, non-dynamical) structures. Isham (1994) distinguishes the following hierarchy of mathematical structures (see also Butterfield and Isham 1999):

Point set of events $\longrightarrow$ topological structure $\longrightarrow$ differentiable manifold $\longrightarrow$ causal structure $\longrightarrow$ Lorentzian structure.

Most approaches subject the Lorentzian and the causal structure to quantization, but keep the manifold structure fixed. More general approaches include quantum topology (Isham 1989), causal sets (see e.g. Sorkin (2005) and Dowker (2006)), and group field theory (see e.g. Oriti 2009). It is assumed therein that space–time is fundamentally discrete; see also the end of Section 6.3 for a brief discussion of such approaches.

According to the Copenhagen interpretation of quantum theory, all structures related to space–time would probably have to stay classical because they are thought to be necessary ingredients for the measurement process; cf. Chapter 10. For the purpose of quantum gravity, such a viewpoint is, however, insufficient and probably inconsistent. The main aim in constructing a quantum theory of gravity is just the opposite: to get rid of any external structure, that is, to implement 'background independence'.

Historically, the first remark on the necessity of dealing with quantum gravity was made by Einstein (1916*b*, p. 108). This was, of course, in the framework of the 'old' quantum theory and does not yet reflect his critical attitude to quantum theory, which he adopted later. After he had recognized that his new theory of general relativity predicted the occurrence of gravitational waves, he wrote[13]:

In the same way, the atoms would have to emit, because of the inner atomic electronic motion, not only electromagnetic, but also gravitational energy, although in tiny amounts. Since this

[13]'Gleichwohl müßten die Atome zufolge der inneratomischen Elektronenbewegung nicht nur elektromagnetische, sondern auch Gravitationsenergie ausstrahlen, wenn auch in winzigem Betrage. Da dies in Wahrheit in der Natur nicht zutreffen dürfte, so scheint es, daß die Quantentheorie nicht nur die Maxwellsche Elektrodynamik, sondern auch die neue Gravitationstheorie wird modifizieren müssen.'

hardly holds true in nature, it seems that quantum theory will have to modify not only Maxwell's electrodynamics, but also the new theory of gravitation.

One of the main motivations for dealing with the problem of quantum gravity was spelled out by Peter Bergmann as follows (Bergmann 1992, p. 364):

Today's theoretical physics is largely built on two giant conceptual structures: quantum theory and general relativity. As the former governs primarily the atomic and subatomic worlds, whereas the latter's principal applications so far have been in astronomy and cosmology, our failure to harmonize quanta and gravitation has not yet stifled progress on either front. Nevertheless, the possibility that there might be some deep dissonance has caused physicists an esthetic unease, and it has caused a number of people to explore avenues that might lead to a quantum theory of gravitation, no matter how many decades away the observations of 'gravitons' might lie in the future.

The following chapters are devoted to the major avenues that are being explored.

Further reading: Ashtekar and Geroch (1974), Carlip (2001), Isham (1994).

2

Covariant approaches to quantum gravity

2.1 The concept of a graviton

A central role in the covariant quantization of the gravitational field is played by the *graviton*—a massless particle of spin 2, which is the mediator of the gravitational interaction. It is analogous to the photon in quantum electrodynamics. Its definition requires, however, the presence of a background structure, at least in an approximate sense. We shall, therefore, first review weak gravitational waves in Minkowski space–time and the concept of helicity. We shall then explain how gravitons are defined as spin-2 particles from representations of the Poincaré group. Finally, the gravitational field in its linear approximation is quantized. We show, in particular, how Poincaré invariance ensures the equivalence principle and therefore leads to the *full* theory of general relativity in the classical limit.

2.1.1 Weak gravitational waves

Our starting point is the decomposition of a space–time metric $g_{\mu\nu}$ into a *fixed* (i.e. non-dynamical) background and a 'perturbation'; see, for example, Weinberg (1972), Misner *et al.* (1973), and Maggiore (2008). In the following, we take for the background the flat Minkowski space–time with the standard metric $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ and call the perturbation $f_{\mu\nu}$. Thus,

$$g_{\mu\nu} = \eta_{\mu\nu} + f_{\mu\nu}. \tag{2.1}$$

We assume that the perturbation is small, that is, that the components of $f_{\mu\nu}$ are small in the standard cartesian coordinates. Using (1.1) without the Λ-term, the Einstein equations (1.3) read in the linear approximation

$$\Box f_{\mu\nu} = -16\pi G \left(T_{\mu\nu} - \tfrac{1}{2}\, \eta_{\mu\nu} T\right), \tag{2.2}$$

where $T := \eta^{\mu\nu} T_{\mu\nu}$, and the 'harmonic condition' (also called the 'de Donder gauge')

$$f_{\mu\nu,}{}^{\nu} = \frac{1}{2} f^{\nu}{}_{\nu,\mu} \tag{2.3}$$

has been used.[1] This condition is analogous to the Lorenz[2] gauge condition in electrodynamics and is used here to partially fix the coordinates. Namely, the invariance of the full theory under coordinate transformations

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x) \tag{2.4}$$

leads to the invariance of the linear theory under (using $\epsilon^2 \approx 0$)

$$f_{\mu\nu} \to f_{\mu\nu} - \epsilon_{\mu,\nu} - \epsilon_{\nu,\mu}. \tag{2.5}$$

Instead of $f_{\mu\nu}$, it is often useful to employ the combination

$$\bar{f}_{\mu\nu} := f_{\mu\nu} - \tfrac{1}{2}\, \eta_{\mu\nu} f^\rho{}_\rho, \tag{2.6}$$

so that (2.2) assumes the simple form

$$\Box \bar{f}_{\mu\nu} = -16\pi G T_{\mu\nu}. \tag{2.7}$$

The harmonic gauge condition (2.3) then reads $\partial_\nu \bar{f}_\mu{}^\nu = 0$, in direct analogy to the Lorenz gauge condition $\partial_\nu A^\nu = 0$. Since (2.7) is analogous to the wave equation $\Box A^\mu = -4\pi j^\mu$, the usual solutions (retarded waves, etc.) can be found. Note that the harmonic gauge condition is consistent with $\partial_\nu T^{\mu\nu} = 0$ (which is analogous to $\partial_\nu j^\nu = 0$), but *not* with $\nabla_\nu T^{\mu\nu} = 0$ (vanishing of covariant derivative). Therefore, although $T_{\mu\nu}$ acts as a source for $f_{\mu\nu}$, there is no exchange of energy in the linear approximation between matter and the gravitational field.

In the vacuum case ($T_{\mu\nu} = 0$), the simplest solutions to (2.2) are plane waves,

$$f_{\mu\nu} = e_{\mu\nu} \mathrm{e}^{\mathrm{i}kx} + e^*_{\mu\nu} \mathrm{e}^{-\mathrm{i}kx}, \tag{2.8}$$

where $e_{\mu\nu}$ is the polarization tensor. One has $k_\mu k^\mu = 0$ and, from (2.3), $k^\nu e_{\mu\nu} = (1/2) k_\mu e^\nu{}_\nu$. With $f_{\mu\nu}$ still obeying (2.3), one can perform a new coordinate transformation of the type (2.4) to get

$$f'_{\mu\nu,}{}^\nu - \tfrac{1}{2}\, f'^\nu{}_{\nu,\mu} = -\Box \epsilon_\mu.$$

Without leaving the harmonic condition (2.3), one can thus *fix* the coordinates by choosing the four functions $\epsilon_\mu(x)$ to satisfy $\Box \epsilon_\mu = 0$ (this equation has plane-wave solutions and is therefore not in conflict with (2.8)). In total, one thus finds $10-4-4 = 2$ independent degrees of freedom for the gravitational field in the linear approximation. The question of how many degrees of freedom the *full* field possesses will be dealt with in Chapter 4. With the ϵ_μ chosen as plane waves, $\epsilon_\mu(x) = 2\,\mathrm{Re}[\mathrm{i} f_\mu \mathrm{e}^{\mathrm{i}kx}]$ (f_μ being real numbers), the transformed plane wave reads the same as (2.8), with

$$e_{\mu\nu} \to e_{\mu\nu} + k_\mu f_\nu + k_\nu f_\mu. \tag{2.9}$$

[1]Indices are raised and lowered by $\eta^{\mu\nu}$ and $\eta_{\mu\nu}$, respectively. We set $c = 1$ in most expressions.

[2]This is not a misprint. The Lorenz condition was first formulated by the Danish physicist Ludvig Valentin Lorenz (1829–91). See Jackson (2008) in this context.

In the case of plane waves, it is most convenient to choose the 'transverse-traceless' (TT) gauge, in which the wave is purely spatial and transverse to its direction of propagation ($e_{\mu\nu}k^\nu = 0$), and $e^\nu{}_\nu = 0$. This turns out to have a gauge-invariant meaning, so that the gravitational waves really are transverse. The two independent linear polarization states are usually called the + polarization and the × polarization (Misner *et al.* 1973).

Consider, for example, a plane wave moving in the $x^1 \equiv x$ direction. In the transverse (y and z) directions, a ring of test particles will be deformed into a pulsating ellipse, with the axis of the + polarization rotated by 45° compared to the × polarization. One has explicitly

$$f_{\mu\nu} = 2\,\mathrm{Re}\left(e_{\mu\nu}\mathrm{e}^{-\mathrm{i}\omega(t-x)}\right), \tag{2.10}$$

with $x^0 \equiv t$, $k^0 = k^1 \equiv \omega > 0$, $k^2 = k^3 = 0$. Denoting by $\mathbf{e}_y$ and $\mathbf{e}_z$ the unit vectors in the y and z directions, respectively, one obtains the expressions

$$e_{22}\mathbf{e}_+ = e_{22}(\mathbf{e}_y \otimes \mathbf{e}_y - \mathbf{e}_z \otimes \mathbf{e}_z) \tag{2.11}$$

and

$$e_{23}\mathbf{e}_\times = e_{23}(\mathbf{e}_y \otimes \mathbf{e}_z + \mathbf{e}_z \otimes \mathbf{e}_y) \tag{2.12}$$

for the polarization tensors for the + and the × polarization, respectively (e_{22} and e_{23} are numbers giving the amplitude of the wave). General solutions of the wave equation can be found by performing *superpositions* of the linear polarization states. In particular,

$$\mathbf{e}_\mathrm{R} = \frac{1}{\sqrt{2}}(\mathbf{e}_+ + \mathrm{i}\mathbf{e}_\times), \quad \mathbf{e}_\mathrm{L} = \frac{1}{\sqrt{2}}(\mathbf{e}_+ - \mathrm{i}\mathbf{e}_\times) \tag{2.13}$$

are the right and the left circular polarization states, respectively. In the general case of elliptic polarization, the ellipse also changes its shape.

Of special interest is the behaviour of the waves under rotation around the axis of propagation (here, the x-axis). Under counterclockwise rotation by an angle θ, the polarization states transform according to

$$\begin{aligned}\mathbf{e}'_+ &= \mathbf{e}_+\cos 2\theta + \mathbf{e}_\times \sin 2\theta,\\ \mathbf{e}'_\times &= \mathbf{e}_\times \cos 2\theta - \mathbf{e}_+ \sin 2\theta.\end{aligned} \tag{2.14}$$

For (2.13), this corresponds to

$$\mathbf{e}'_\mathrm{R} = \mathrm{e}^{-2\mathrm{i}\theta}\mathbf{e}_\mathrm{R}\,, \quad \mathbf{e}'_\mathrm{L} = \mathrm{e}^{2\mathrm{i}\theta}\mathbf{e}_\mathrm{L}. \tag{2.15}$$

The polarization tensors thus rotate through an angle 2θ. This corresponds to a symmetry with respect to a rotation by 180°.

If a plane wave φ transforms as $\varphi \to \mathrm{e}^{\mathrm{i}h\theta}$ under a rotation around the direction of propagation, one calls h its *helicity*. The left and right circularly polarized gravitational waves thus have helicity 2 and -2, respectively. In the quantum theory, these states will become the states of the 'graviton'; see Section 2.1.2. For plane waves with helicity

h, the angle of mutual inclination of the axes of linear polarization is $90°/h$. For a spin-1/2 particle, for example, this is 180°, which is why invariance for them is only reached after a rotation by 720°.

For electrodynamics, the right and left polarized states are given by the *vectors*

$$\mathbf{e}_{\mathrm{R}} = \frac{1}{\sqrt{2}}(\mathbf{e}_y + \mathrm{i}\mathbf{e}_z)\ , \quad \mathbf{e}_{\mathrm{L}} = \frac{1}{\sqrt{2}}(\mathbf{e}_y - \mathrm{i}\mathbf{e}_z). \tag{2.16}$$

Under the above rotation, they transform as

$$\mathbf{e}'_{\mathrm{R}} = \mathrm{e}^{-\mathrm{i}\theta}\mathbf{e}_{\mathrm{R}}\ , \quad \mathbf{e}'_{\mathrm{L}} = \mathrm{e}^{\mathrm{i}\theta}\mathbf{e}_{\mathrm{L}}. \tag{2.17}$$

The left and right circularly polarized electromagnetic waves thus have helicity 1 and -1, respectively. Instead of (2.8), one has in this case

$$A_\mu = e_\mu \mathrm{e}^{\mathrm{i}kx} + e^*_\mu \mathrm{e}^{-\mathrm{i}kx} \tag{2.18}$$

with $k_\mu k^\mu = 0$ and $k_\nu e^\nu = 0$. It is possible to perform a gauge transformation without leaving the Lorenz gauge, $A_\mu \to A'_\mu = A_\mu + \partial_\mu\Lambda$ with $\Box\Lambda = 0$. If Λ is a plane-wave solution, $\Lambda = 2\,\mathrm{Re}[\mathrm{i}\lambda \mathrm{e}^{\mathrm{i}kx}]$, one has instead of (2.9)

$$e_\mu \to e_\mu - \lambda k_\mu. \tag{2.19}$$

The field equations of linearized gravity can be obtained from the following Lagrangian (Fierz and Pauli 1939):

$$\begin{aligned}\mathcal{L} = \frac{1}{64\pi G}\ & \left(f^{\mu\nu,\sigma} f_{\mu\nu,\sigma} - f^{\mu\nu,\sigma} f_{\sigma\nu,\mu} - f^{\nu\mu,\sigma} f_{\sigma\mu,\nu}\right. \\ & \left. - f^{\mu}{}_{\mu,\nu} f^{\rho}{}_{\rho,}{}^{\nu} + 2 f^{\rho\nu}{}_{,\nu} f^{\sigma}{}_{\sigma,\rho}\right) - \frac{1}{2} T_{\mu\nu} f^{\mu\nu}.\end{aligned} \tag{2.20}$$

The Euler–Lagrange field equations are (writing $f := f^{\mu}{}_{\mu}$)

$$f_{\mu\nu,\sigma}{}^{\sigma} - f_{\sigma\mu,\nu}{}^{\sigma} - f_{\sigma\nu,\mu}{}^{\sigma} + f_{,\mu\nu} + \eta_{\mu\nu}\left(f^{\alpha\beta}{}_{,\alpha\beta} - f_{,\sigma}{}^{\sigma}\right) = -16\pi G T_{\mu\nu}. \tag{2.21}$$

The left-hand side of this equation is -2 times the Einstein tensor, $-2G_{\mu\nu}$, in the linear approximation. It obeys the linearized Bianchi identity $\partial_\nu G^{\mu\nu} = 0$, which is consistent with $\partial_\nu T^{\mu\nu} = 0$. The Bianchi identity is a consequence of the gauge invariance (modulo a total divergence) of the Lagrangian (2.20) with respect to (2.4).[3]

Taking in (2.21) the trace and substituting the $\eta_{\mu\nu}$ term, one gets

$$\Box f_{\mu\nu} - f_{\sigma\mu,\nu}{}^{\sigma} - f_{\sigma\nu,\mu}{}^{\sigma} + f_{,\mu\nu} = -16\pi G\left(T_{\mu\nu} - \tfrac{1}{2}\,\eta_{\mu\nu} T\right). \tag{2.22}$$

Using the harmonic condition (2.3), one finds the linearized Einstein equations (2.2).[4]

[3] The Bianchi identity in electrodynamics reads $\partial_\mu(\partial_\nu F^{\mu\nu}) = 0$, consistent with the charge-conservation law $\partial_\nu j^\nu = 0$.

[4] It is often useful to make the redefinition $f_{\mu\nu} \to \sqrt{32\pi G} f_{\mu\nu}$; cf. Section 2.2.2.

Exploiting the Poincaré invariance of the flat background, one can calculate from the Fierz–Pauli Lagrangian (2.20) without the $T_{\mu\nu}$ term the *canonical* energy–momentum tensor of the linearized gravitational field,

$$t_{\mu\nu} := \frac{\partial \mathcal{L}}{\partial f_{\alpha\beta,}{}^{\nu}} f_{\alpha\beta,\mu} - \eta_{\mu\nu}\mathcal{L}. \tag{2.23}$$

The resulting expression is lengthy, but can be considerably simplified in the TT gauge, in which $f_{\mu\nu}$ assumes the form

$$f_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & f_{22} & f_{23} \\ 0 & 0 & f_{23} & -f_{22} \end{pmatrix}.$$

Then,

$$t_{00} \stackrel{\mathrm{TT}}{=} \frac{1}{16\pi G}\left(\dot{f}_{22}^2 + \dot{f}_{23}^2\right) = -t_{01}, \tag{2.24}$$

which can be written covariantly as

$$t_{\mu\nu} \stackrel{\mathrm{TT}}{=} \frac{1}{32\pi G} f_{\alpha\beta,\mu} f^{\alpha\beta}{}_{,\nu}. \tag{2.25}$$

It is sometimes appropriate to average this expression over a region of space–time much larger than ω^{-1}, where ω is the frequency of the weak gravitational wave, so that terms such as $\exp(-2\mathrm{i}\omega(t-x))$ drop out. In the TT gauge, this leads to

$$\bar{t}_{\mu\nu} = \frac{k_\mu k_\nu}{16\pi G} e^{\alpha\beta *} e_{\alpha\beta}. \tag{2.26}$$

In the general harmonic gauge, that is, without particularization to the TT gauge, the term $e^{\alpha\beta *}e_{\alpha\beta}$ is replaced by $e^{\alpha\beta *}e_{\alpha\beta} - \frac{1}{2}|e^{\alpha}{}_{\alpha}|^2$. This expression remains invariant under the gauge transformations (2.9), as it should.

Regarding the Fierz–Pauli Lagrangian (2.20), the question arises whether it could serve as a candidate for a fundamental helicity-2 theory of the gravitational field *in* a flat background. As it stands, this is certainly not possible because, as already mentioned, one has $\partial_\nu T^{\mu\nu} = 0$, and there is therefore no back reaction of the gravitational field on the matter. One might thus wish to add the canonical energy–momentum tensor $t_{\mu\nu}$ (2.23) to the right-hand side of the linearized Einstein equations:

$$\Box \bar{f}_{\mu\nu} = -16\pi G(T_{\mu\nu} + t_{\mu\nu}).$$

This modified equation would, however, lead to a Lagrangian *cubic* in the fields, which in turn would give a new contribution to $t_{\mu\nu}$, and so on. Deser (1970) was able to show that this infinite process can actually be performed in one single step; see also Deser (2010). The result is that the original metric $\eta_{\mu\nu}$ is unobservable and that all matter couples to the metric $g_{\mu\nu} = \eta_{\mu\nu} + f_{\mu\nu}$; the resulting action is the Einstein–Hilbert action, and the theory is therefore GR. Minkowski space–time as a background structure has completely disappeared. Boulanger *et al.* (2001) have shown that, starting from a finite number of Fierz–Pauli Lagrangians, no consistent coupling between the various helicity-2 fields is possible if the fields occur at most with second derivatives—one merely obtains a sum of uncoupled Einstein–Hilbert actions.

Since GR follows uniquely from (2.20), the question arises whether one would be able to construct a pure scalar, fermionic, or vector theory of gravity; cf. Feynman *et al.* (1995) and Straumann (2000). As Maxwell already knew, a vector theory is excluded because it would lead to repulsive forces. Fermions are excluded because an object that emits a fermion does not remain in the same internal state (there are also problems with a two-fermion exchange). A scalar theory, in contrast, would lead to attraction. In fact, even before the advent of GR, Nordström had tried to describe gravity by a scalar theory, which can be defined by the Lagrangian

$$\mathcal{L} = -\frac{1}{2}\,\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - 4\pi GT\phi + \mathcal{L}_{\mathrm{matter}}, \tag{2.27}$$

where $T = \eta_{\mu\nu}T^{\mu\nu}$. This leads to the field equation

$$\Box\phi = 4\pi GT. \tag{2.28}$$

The physical metric (as measured by rods and clocks) turns out to be

$$g_{\mu\nu}(x) \equiv \phi^2(x)\eta_{\mu\nu}.$$

It is thus conformally flat and possesses a vanishing Weyl tensor. A non-linear generalization of the Nordström theory was given by Einstein and Fokker (1914); their field equations read

$$R = 24\pi GT. \tag{2.29}$$

However, this theory is in contradiction with observation, since it does not implement an interaction between gravity and the electromagnetic field (the latter has $T = 0$), and the perihelion motion of Mercury comes out incorrectly. Moreover, this theory contains an absolute structure (cf. Section 1.3): the conformal structure (the 'lightcone') is given from the outset and the theory thus possesses an invariance group (the conformal group), which in four dimensions is a finite-dimensional Lie group and which must be conceptually distinguished from the diffeomorphism group of GR. While pure scalar fields are thus unsuitable for a theory of the gravitational field, they can nevertheless occur *in addition* to the metric of GR. In fact, this happens quite frequently in unified theories; cf. Chapter 9.

2.1.2 Gravitons from representations of the Poincaré group

We shall now turn to the quantum theory of the linear gravitational field. The discussion in the previous subsection suggests that it is described by the behaviour of a massless spin-2 particle. Why massless? From the long-range nature of the gravitational interaction, it is clear that the graviton must have a small mass. Since the presence of a non-vanishing mass, however small, affects the deflection of light discontinuously, one may argue that the graviton mass is strictly zero; see van Dam and Veltman (1970) and Zakharov (1970). These authors have shown that the massive version of the Fierz–Pauli Lagrangian (2.20) does not lead to linearized GR in the limit when the mass is set to zero.[5] A similar argument cannot be put forward for the photon.

In the following, we shall give a brief derivation of the spin-2 nature in the framework of representation theory (see e.g. Weinberg 1995 or Sexl and Urbantke 2001). In this subsection, we shall only deal with one-particle states ('quantum mechanics'), while field-theoretic aspects will be discussed in Section 2.1.3. The important ingredient is the presence of flat Minkowski space–time with metric $\eta_{\mu\nu}$ as an absolute background structure, and the ensuing Poincaré symmetry. The use of the *Poincaré group*, not available beyond the linearized level, already indicates the approximate nature of the graviton concept. We shall set $\hbar = 1$ in most of the following expressions.

According to Wigner, 'particles' are classified by irreducible representations of the Poincaré group. We describe a Poincaré transformation as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \tag{2.30}$$

where the $\Lambda^\mu{}_\nu$ denote Lorentz transformations and the a^μ denote space–time translations. According to Wigner's theorem, (2.30) induces a unitary transformation[6] in the *Hilbert space* of the theory,

$$\psi \to U(\Lambda, a)\psi. \tag{2.31}$$

This ensures that probabilities remain unchanged under the action of the Poincaré group. Since this is a Lie group, it is advantageous to study group elements close to the identity,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu \ , \quad a^\mu = \epsilon^\mu, \tag{2.32}$$

where $\omega_{\mu\nu} = -\omega_{\nu\mu}$. This corresponds to the unitary transformation[7]

$$U(1+\omega, \epsilon) = 1 + \tfrac{1}{2}\,\mathrm{i}\,\omega_{\mu\nu}J^{\mu\nu} - \mathrm{i}\,\epsilon_\mu P^\mu + \dots \ , \tag{2.33}$$

where $J^{\mu\nu}$ and P^μ denote the 10 Hermitian generators of the Poincaré group, which are the boost generators, the angular momentum and the four-momentum. They obey the Lie-algebra relations

$$[P^\mu, P^\rho] = 0 \ , \tag{2.34}$$
$$\mathrm{i}[J^{\mu\nu}, J^{\lambda\rho}] = \eta^{\nu\lambda}J^{\mu\rho} - \eta^{\mu\lambda}J^{\nu\rho} - \eta^{\rho\mu}J^{\lambda\nu} + \eta^{\rho\nu}J^{\lambda\mu} \ , \tag{2.35}$$
$$\mathrm{i}[P^\mu, J^{\lambda\rho}] = \eta^{\mu\lambda}P^\rho - \eta^{\mu\rho}P^\lambda. \tag{2.36}$$

One-particle states are classified according to their behaviour with respect to Poincaré transformations. Since the components P^μ of the four-momentum commute with each other, we shall choose their eigenstates,

[5]It is, however, possible that the limit of GR may be obtained from a non-linear Fierz–Pauli Lagrangian via the so-called Vainshtein mechanism, see Babichev *et al.* (2010) in this context.

[6]The theorem also allows anti-unitary transformations, but these are relevant only for discrete symmetries.

[7]The plus sign on the right-hand side is enforced by the commutation relations $[J_i, J_k] = \mathrm{i}\epsilon_{ikl}J_l$ (where $J_3 \equiv J_{12}$, etc.), whereas the minus sign in front of the second term is pure convention.

$$P^\mu \psi_{p,\sigma} = p^\mu \psi_{p,\sigma}, \tag{2.37}$$

where σ stands symbolically for all other variables. Application of the unitary operator then yields

$$U(1,a)\psi_{p,\sigma} = \mathrm{e}^{-\mathrm{i}p^\mu a_\mu}\psi_{p,\sigma}. \tag{2.38}$$

How do these states transform under Lorentz transformations (we only consider orthochronous proper transformations)? According to the method of induced representations, it is sufficient to find the representations of the *little group*. This group is characterized by the fact that it leaves a 'standard' vector k^μ invariant (within each class of given $p^2 \leq 0$[8] and given sign of p^0). For positive p^0, one can distinguish the following two cases. The first possibility is $p^2 = -m^2 < 0$. Here one can choose $k^\mu = (m, 0, 0, 0)$, and the little group is SO(3), since these are the only Lorentz transformations that leave a particle with $\mathbf{k} = 0$ at rest. The second possibility is $p^2 = 0$. One chooses $k^\mu = (1, 0, 0, 1)$, and the little group is ISO(2), the invariance group of Euclidean geometry (rotations and translations in two dimensions). Any p^μ within a given class can be obtained from the corresponding k^μ by a Lorentz transformation. The normalization is chosen such that

$$\langle \psi_{p',\sigma'}, \psi_{p,\sigma}\rangle = \delta_{\sigma\sigma'}\delta(\mathbf{p}-\mathbf{p}'). \tag{2.39}$$

Consider first the case $m \neq 0$, for which the little group is SO(3). As is well known from quantum mechanics, its unitary representations are a direct sum of irreducible unitary representations $\mathcal{D}^{(j)}_{\sigma\sigma'}$ with dimensions $2j+1$ ($j = 0, \frac{1}{2}, 1, \ldots$). Denoting the angular momentum with respect to the z-axis by $J_3^{(j)} \equiv J_{12}^{(j)}$, one obtains $(J_3^{(j)})_{\sigma\sigma'} = \sigma\delta_{\sigma\sigma'}$ with $\sigma = -j, \ldots, +j$.

In contrast, for $m = 0$, the little group is ISO(2). This is the case of interest here. It turns out that the quantum-mechanical states are only distinguished by the eigenvalue of J_3, the component of the angular momentum in the direction of motion (recall $k^\mu = (1, 0, 0, 1)$ from above):

$$J_3\psi_{k,\sigma} = \sigma\psi_{k,\sigma}. \tag{2.40}$$

The eigenvalue σ is called the *helicity*. One then gets

$$U(\Lambda, 0)\psi_{p,\sigma} = N\mathrm{e}^{\mathrm{i}\sigma\theta(\Lambda,p)}\psi_{\Lambda p,\sigma}, \tag{2.41}$$

where θ denotes the angle in the rotation part of Λ. Since massless particles are not at rest in any inertial system, helicity is a Lorentz-invariant property and may be used to characterize a particle with $m = 0$. There remain only the two possibilities of having the angular momentum parallel or opposite to the direction of propagation.

Comparison with (2.17) shows that the relation $\sigma = \pm 1$ characterizes the *photon*.[9] Because weak gravitational waves in a flat background have helicity 2 (see (2.15)), we associate the particle with $\sigma = \pm 2$ with the gravitational interaction and call it

[8] This restriction is imposed in order to avoid tachyons (particles with $m^2 < 0$).

[9] Because of space inversion symmetry, $\sigma = 1$ and $\sigma = -1$ describe the same particle. This holds also for the graviton, but due to parity violation not, for example, for (massless) neutrinos.

the *graviton*. Since for a massless particle $|\sigma|$ is called its *spin*, we recognize that the graviton has spin 2. The helicity eigenstates (2.41) correspond to circular polarization (see Section 2.1.1), while their superpositions correspond in the generic case to elliptic polarization, or (for equal absolute values of the amplitudes) to linear polarization.

When talking about a massless graviton, an interesting question is, of course, how strong the experimental limits for a possible mass $m_{\rm g}$ are. Goldhaber and Nieto (2010) give an overview of existing limits. The strongest limit comes from gravitational effects on the scale of galaxy clusters, which gives $m_{\rm g} \lesssim 10^{-29}$ eV; this mass corresponds to a (reduced) Compton wavelength of 2×10^{22} m. This limit was found from a Yukawa modification to the potential at long distances for quasi-static gravitational fields. For non-static fields, the experimental limits are weaker; from the absence of gravitational-wave dispersion, for example, one gets $m_{\rm g} \lesssim 8 \times 10^{-20}$ eV.[10]

2.1.3 Quantization of the linear field theory

We now turn to field theory. One starts from a superposition of plane-wave solutions (2.8) and formally turns it into an operator,

$$f_{\mu\nu}(x) = \sum_{\sigma=\pm 2} \int \frac{{\rm d}^3 k}{\sqrt{2|\mathbf{k}|}} \left[a(\mathbf{k},\sigma) e_{\mu\nu}(\mathbf{k},\sigma) {\rm e}^{{\rm i}kx} + a^\dagger(\mathbf{k},\sigma) e^*_{\mu\nu}(\mathbf{k},\sigma) {\rm e}^{-{\rm i}kx} \right]. \tag{2.42}$$

As in the usual interpretation of free quantum field theory, $a(\mathbf{k},\sigma)$ and $a^\dagger(\mathbf{k},\sigma)$ are interpreted, respectively, as the annihilation and creation operators for a graviton of momentum $\hbar\mathbf{k}$ and helicity σ (see e.g. Weinberg 1995). They obey

$$[a(\mathbf{k},\sigma), a^\dagger(\mathbf{k}',\sigma')] = \delta_{\sigma\sigma'} \delta(\mathbf{k} - \mathbf{k}'), \tag{2.43}$$

with all other commutators vanishing. The quantization of the linearized gravitational field was discussed by Bronstein (1936) and Gupta (1952*a*). The latter applied the quantization method that he had developed earlier for electrodynamics, which became known as the 'Gupta–Bleuler method'. It makes use of an indefinite metric in Fock space. He could show in an explicit manner that only two degrees of freedom remain, corresponding to the two polarization states of the massless graviton with spin 2.

Since we only want the presence of helicities ± 2, $f_{\mu\nu}$ cannot be a true tensor with respect to Lorentz transformations (note that the TT gauge condition is not Lorentz invariant). As a consequence, one is *forced* to introduce gauge invariance and demand that $f_{\mu\nu}$ transform under a Lorentz transformation according to

$$f_{\mu\nu} \to \Lambda_\mu{}^\lambda \Lambda_\nu{}^\rho f_{\lambda\rho} - \partial_\nu \epsilon_\mu - \partial_\mu \epsilon_\nu, \tag{2.44}$$

in order to stay within the TT gauge; cf. (2.5). Therefore, the coupling in the Lagrangian (2.20) must be to a *conserved* source, $\partial_\nu T^{\mu\nu} = 0$, because otherwise the coupling is not gauge invariant.

The occurrence of gauge invariance can also be understood in a group-theoretic way. We start with a symmetric tensor field $f_{\mu\nu}$ with vanishing trace (nine degrees of

[10] For comparison, the strongest experimental limit on the photon mass is $m_\gamma \lesssim 10^{-18}$ eV (Goldhaber and Nieto 2010).

freedom). This field transforms according to the irreducible $\mathcal{D}^{(1,1)}$ representation of the Lorentz group; see, for example, Section 5.6 in Weinberg (1995). Its restriction to the subgroup of rotations yields

$$\mathcal{D}^{(1,1)} = \mathcal{D}^{(1)} \otimes \mathcal{D}^{(1)} , \tag{2.45}$$

where $\mathcal{D}^{(1)}$ denotes the $j = 1$ representation of the rotation group. (Since one has three 'angles' in each $\mathcal{D}^{(1)}$, this yields the $3 \times 3 = 9$ degrees of freedom of the trace-free $f_{\mu\nu}$.) The representation (2.45) is reducible with the 'Clebsch–Gordan decomposition'

$$\mathcal{D}^{(1)} \otimes \mathcal{D}^{(1)} = \mathcal{D}^{(2)} \oplus \mathcal{D}^{(1)} \oplus \mathcal{D}^{(0)}, \tag{2.46}$$

corresponding to the five degrees of freedom for a massive spin-2 particle, three degrees of freedom for a spin-1 particle, and one degree of freedom for a spin-0 particle, respectively. The $3+1$ degrees of freedom for spin 1 and spin 0 can be eliminated by the four conditions $\partial_\nu f^{\mu\nu} = 0$ (transversality). To obtain only two degrees of freedom one needs to take into account again the gauge freedom (2.5). This yields three additional conditions, since the four ϵ_μ in (2.5) have to satisfy the condition $\partial_\mu \epsilon^\mu = 0$ in order to preserve tracelessness. (The condition $\Box\epsilon_\mu = 0$—needed to preserve transversality—is automatically fulfilled for plane waves). In total, one arrives at $(10-1)-4-3 = 2$ degrees of freedom, corresponding to the two helicity states of the graviton.[11]

The same arguments also apply of course to electrodynamics: A^μ cannot transform as a Lorentz vector, since, for example, the temporal gauge $A^0 = 0$ can be chosen. Instead, one is forced to introduce gauge invariance, and the transformation law is

$$A_\mu \to \Lambda_\mu^{\ \nu} A_\nu + \partial_\mu \epsilon, \tag{2.47}$$

in analogy to (2.44). Therefore, a Lagrangian is needed that couples to a *conserved* source, $\partial_\mu j^\mu = 0$.[12] The group-theoretic argument for QED goes as follows. A vector field transforms according to the $\mathcal{D}^{(1/2,1/2)}$ representation of the Lorentz group, which, if restricted to rotations, can be decomposed as

$$\mathcal{D}^{(1/2,1/2)} = \mathcal{D}^{(1/2)} \otimes \mathcal{D}^{(1/2)} = \mathcal{D}^{(1)} \oplus \mathcal{D}^{(0)}. \tag{2.48}$$

The $\mathcal{D}^{(0)}$ describes spin 0, which is eliminated by the Lorenz condition $\partial_\nu A^\nu = 0$. The $\mathcal{D}^{(1)}$ corresponds to the three degrees of freedom of a massive spin-1 particle. One of these is eliminated by the gauge transformation $A_\mu \to A_\mu + \partial_\mu \epsilon$ (with $\Box\epsilon = 0$ to preserve the Lorenz condition) to arrive at the two degrees of freedom for the massless photon.

Weinberg (1964) concluded (see also Weinberg 1995, p. 537) that one can derive the equivalence principle (and thus GR if no other fields are present) from the *Lorentz invariance* of the spin-2 theory (plus the pole structure of the S-matrix). Similar arguments can be put forward in the electromagnetic case to show that electric charge

[11] The counting in the canonical version of the theory leads of course to the same result and is presented in Section 4.2.3.

[12] This is also connected with the fact that massless spins ≥ 3 are usually excluded, because no conserved tensor is available. Massless spins ≥ 3 cannot generate long-range forces; cf. Weinberg (1995, p. 252).

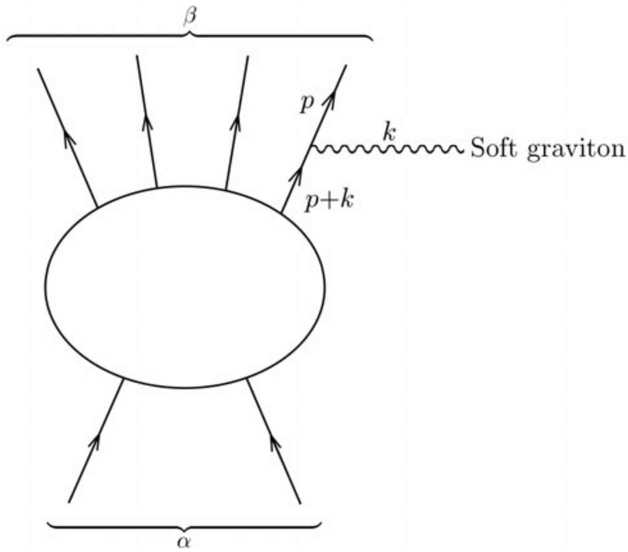

Fig. 2.1 Emission of a soft graviton.

must be conserved. No gauge invariance arguments are needed, at least not explicitly. The gravitational mass $m_{\rm g}$ is defined in this approach by the strength of interaction with a soft graviton, that is, a graviton with four-momentum $k \to 0$. The amplitude for the emission (see Fig. 2.1) of a single soft graviton is given by

$$M^{\mu\nu}_{\beta\alpha}(k) = M_{\beta\alpha} \cdot \sum_n \frac{\eta_n g_n p^\mu_n p^\nu_n}{p^\mu_n k_\mu - \mathrm{i}\eta_n \epsilon}, \tag{2.49}$$

where $M_{\beta\alpha}$ denotes the amplitude for the process without soft-graviton emission, and the sum runs over all ingoing and outgoing particles; α refers to the ingoing and β to the outgoing particles (Fig. 2.1); g_n denotes the coupling of the graviton to particle n ($\eta_n = 1$ for outgoing and $\eta_n = -1$ for ingoing particles); and p_n is the four-momentum of the nth particle (in the initial and final state, respectively).

To calculate the amplitude for graviton emission with definite helicity, one must contract (2.49) with the polarization tensor $e_{\mu\nu}(\mathbf{k}, \sigma)$. As argued above, however, the latter is not a Lorentz tensor, but transforms according to

$$e_{\mu\nu} \to \Lambda_\mu{}^\lambda \Lambda_\nu{}^\rho e_{\lambda\rho} + k_\nu f_\mu + k_\mu f_\nu .$$

To obtain a Lorentz-invariant amplitude, one must therefore demand that

$$k_\mu M^{\mu\nu}_{\beta\alpha}(k) = 0.$$

From (2.49), one then finds

$$\sum_n \eta_n g_n p_n^\nu = 0,$$

which is equivalent to the statement that $\sum_n g_n p_n^\nu$ is conserved. But, for non-trivial processes, the only linear combination of momenta that is conserved is the total momentum $\sum_n \eta_n p_n^\nu$. Consequently, the couplings g_n must all be *equal*, and one can set $g_n \equiv \sqrt{8\pi G}$. Therefore, all low-energy particles with spin 2 and $m = 0$ couple to all forms of energy in an equal way. As in Section 2.1.1, this shows the 'equivalence' of a spin-2 theory with GR. From this point of view, GR is a consequence of quantum theory. Weinberg (1964) also showed that the effective gravitational mass $m_{\rm g}$ is given by

$$m_{\rm g} = 2E - \frac{m_{\rm i}^2}{E},$$

where, as in Section 1.1.4, $m_{\rm i}$ denotes the inertial mass, and E is the energy. For $E \to m_{\rm i}$, this leads to the usual equivalence of inertial and gravitational mass. At the same time, one has $m_{\rm g} = 2E$ for $m_{\rm i} \to 0$. How can this be interpreted? In his 1911 calculation of the deflection of light, Einstein found the Newtonian expression for the deflection angle from the equivalence principle alone (setting $m_{\rm g} = E$). The full theory of GR, however, yields twice this value, corresponding to $m_{\rm g} = 2E$.

Weinberg's arguments, as well as the approaches presented in Section 2.1.1, are important for unified theories such as string theory (see Chapter 9), in which a massless spin-2 particle emerges with necessity, leading to the claim that such theories contain GR in an appropriate limit.

In this discussion of linear quantum gravity, the question arises whether this framework leads to observable effects in the laboratory. A brief estimate suggests that such effects would be too tiny. In atomic physics, we can compare the quantum-gravitational decay rate $\Gamma_{\rm g}$ with its electromagnetic counterpart $\Gamma_{\rm e}$; we expect for dimensional reasons

$$\frac{\Gamma_{\rm g}}{\Gamma_{\rm e}} \sim \alpha^n \left(\frac{m_{\rm e}}{m_{\rm P}}\right)^2 \tag{2.50}$$

with some power n of the fine-structure constant α, and where $m_{\rm e}$ is the electron mass. The square of the mass ratio already yields the tiny number 10^{-45}. Still, it is instructive to discuss some examples in detail. Following Weinberg (1972), one can calculate the transition rate from the $3d$ level to the $1s$ level in the hydrogen atom due to the emission of a graviton. One needs at least the $3d$ level, since $\Delta l = 2$ is needed for the emission of a spin-2 particle.

One starts from the classical formula for gravitational radiation and interprets it as the emission rate of gravitons with energy $\hbar\omega$,

$$\Gamma_{\rm g} = \frac{P}{\hbar\omega}, \tag{2.51}$$

where P denotes the classical expression for the emitted power. This was already suggested by Bronstein (1936). In the quadrupole approximation, one then finds for the transition rate from an initial state i to a final state f (restoring c)

$$\Gamma_{\rm g}(i \to f) = \frac{2G\omega^5}{5\hbar c^5}\left(\sum_{kl} Q^*_{kl}(i \to f) Q_{kl}(i \to f) - \frac{1}{3}\sum_k |Q_{kk}(i \to f)|^2\right), \tag{2.52}$$

where

$$Q_{kl}(i \to f) := m_{\rm e} \int {\rm d}^3 x \; \psi_f^*(\mathbf{x}) x_k x_l \psi_i(\mathbf{x}) \tag{2.53}$$

is the quantum-mechanical analogue of the classical quadrupole moment,

$$Q_{kl} = \frac{1}{c^2}\int {\rm d}^3x \; x_k x_l T_{00}. \tag{2.54}$$

Inserting the hydrogen eigenfunctions ψ_{100} and ψ_{32m} $(m = -2, \ldots, 2)$ for ψ_i and ψ_f, respectively, and averaging over m, one finds after some calculations[13]

$$\Gamma_{\rm g} = \frac{G m_{\rm e}^3 c \alpha^6}{360\hbar^2} \approx 5.7 \times 10^{-40} \ {\rm s}^{-1}. \tag{2.55}$$

This corresponds to a lifetime of

$$\tau_{\rm g} \approx 5.6 \times 10^{31} \ \text{years}, \tag{2.56}$$

which is too long to be observable (although it is of the same order as the value for the proton lifetime predicted by some unified theories). Since electromagnetic dipole transitions are of the order of $\Gamma_{\rm e} \sim m_{\rm e}\alpha^5 c^2/\hbar$, one has

$$\frac{\Gamma_{\rm g}}{\Gamma_{\rm e}} \sim \alpha \left(\frac{m_{\rm e}}{m_{\rm P}}\right)^2 \approx 1.28 \times 10^{-47} \; ; \tag{2.57}$$

cf. (2.50). The result (2.55) is in fact much more general than its derivation; field-theoretic methods lead to exactly the same decay rate (Boughn and Rothman 2006).

Through similar heuristic considerations, one can also calculate the number N of gravitons per unit volume associated with a gravitational wave (Weinberg 1972). We recall from (2.26) and the accompanying remarks that the space–time average of the energy–momentum tensor for weak gravitational waves is given in the general harmonic gauge by the expression

$$\bar{t}_{\mu\nu} = \frac{k_\mu k_\nu c^4}{16\pi G}\left(e^{\alpha\beta *} e_{\alpha\beta} - \frac{1}{2}|e^{\alpha}{}_{\alpha}|^2\right). \tag{2.58}$$

Since the energy–momentum tensor of a collection of particles is (cf. Weinberg 1972, Section 2.8)

$$T^{\mu\nu} = c^2 \sum_n \frac{p_n^\mu p_n^\nu}{E_n} \delta(\mathbf{x} - \mathbf{x}_n(t)), \tag{2.59}$$

[13] The result is different from Weinberg (1972) because it seems that Weinberg used the eigenfunctions from Messiah's textbook on quantum mechanics, which contain a misprint. I am grateful to N. Straumann for discussions on this point.

a collection of gravitons with momenta $p^\mu = \hbar k^\mu$ possesses the energy–momentum tensor

$$T_{\rm g}^{\mu\nu} = \frac{\hbar c^2 k^\mu k^\nu N}{\omega}. \tag{2.60}$$

Comparison with (2.58) then yields

$$N = \frac{\omega c^2}{16\pi\hbar G}\left(e^{\alpha\beta *}e_{\alpha\beta} - \frac{1}{2}|e^\alpha{}_\alpha|^2\right), \tag{2.61}$$

where the expression in parentheses is the amplitude squared, $|A|^2$. Taking, for example, $\omega \approx 1$ kHz and $|A| \approx 10^{-21}$ (typical values for gravitational waves arriving from a supernova), one finds $N \approx 3 \times 10^{14}$ cm^{-3}. For the stochastic background of gravitons from the early universe, one would expect typical values of $\omega \approx 10^{-3}$ Hz and $|A| \approx 10^{-22}$, leading to $N \approx 10^6$ cm^{-3}. Effects of single gravitons are thus not expected to be observable in gravitational-wave experiments, at least not in the near future (one would have to reduce $\omega|A|^2$ by another factor of 10^6).

The ground-state wave function of a quantum-mechanical oscillator is given by a Gaussian, $\psi_0(x) \propto \exp(-\omega m x^2/2\hbar)$. Similarly, one can express the ground state of a free quantum field as a Gaussian wave *functional*. We shall discuss such functionals in detail in Section 5.3, but mention here that there exists an explicit form for the ground-state functional of the linear graviton field discussed above (Kuchař 1970). This has a form similar to the ground-state functional of free QED, which reads

$$\Psi_0[\mathbf{A}] \propto \exp\left(-\frac{1}{2}\int \mathrm{d}^3x\ \mathrm{d}^3y\ A^a(\mathbf{x})\omega_{ab}(\mathbf{x},\mathbf{y})A^b(\mathbf{y})\right) \tag{2.62}$$

with

$$\omega_{ab}(\mathbf{x},y) = (-\nabla^2\delta_{ab} + \partial_a\partial_b)\int \mathrm{d}^3p\ \frac{\mathrm{e}^{-\mathrm{i}\mathbf{p}(\mathbf{x}-\mathbf{y})}}{|\mathbf{p}|}. \tag{2.63}$$

Note that a and b are only spatial indices, since the wave functional is defined on space-like hypersurfaces. This state can also be written in a manifestly gauge-invariant form,

$$\Psi_0 \propto \exp\left(-\frac{c}{4\pi^2\hbar}\int \mathrm{d}^3x\ \mathrm{d}^3y\ \frac{\mathbf{B}(\mathbf{x})\mathbf{B}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^2}\right), \tag{2.64}$$

where $\mathbf{B}$ is the magnetic field, and c and $\hbar$ have been reinserted. In the functional picture, it is evident that the ground state is a highly non-local state. The ground state of the linearized gravitational field can be written in a manifestly gauge-invariant way similar to (2.64) as

$$\Psi_0 \propto \exp\left(-\frac{c^3}{8\pi^2 G\hbar}\int \mathrm{d}^3x\ \mathrm{d}^3y\ \frac{f^{\rm TT}_{ab,c}(\mathbf{x})f^{\rm TT}_{ab,c}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^2}\right), \tag{2.65}$$

where f_{ab} are the spatial components of the metric perturbation introduced in (2.1).

2.2 Path-integral quantization

2.2.1 General properties of path integrals

A popular method in both quantum mechanics and quantum field theory is path-integral quantization (Feynman and Hibbs 1965). In the case of quantum mechanics, the propagator for a 'particle' to go from position $\mathbf{x}'$ at time t' to position $\mathbf{x}''$ at time t'' can be expressed as a formal sum over all possible paths connecting these positions,

$$\langle \mathbf{x}'', t'' | \mathbf{x}', t' \rangle = \int \mathcal{D}\mathbf{x}(t)\ \mathrm{e}^{\mathrm{i}S[\mathbf{x}(t)]/\hbar}. \tag{2.66}$$

It is important to remember that most 'paths' in this sum are continuous, but nowhere differentiable, and that $\mathcal{D}\mathbf{x}(t)$ is, in fact, a formal notation for the following limiting process (spelled out for simplicity for a particle in one spatial dimension):

$$\begin{aligned}\langle x'', t'' | x', t' \rangle = \lim_{N\to\infty} \int \mathrm{d}x_1 \cdots \mathrm{d}x_{N-1} \left(\frac{mN}{2\pi \mathrm{i} t\hbar}\right)^{N/2} \\ \times \prod_{j=0}^{N-1} \exp\left(-\frac{m(x_{j+1}-x_j)^2 N}{2\mathrm{i}t\hbar} - \frac{\mathrm{i}tV(x_j)}{\hbar N}\right),\end{aligned} \tag{2.67}$$

where m denotes the mass, V the potential, and $t \equiv t'' - t'$. The path integral is especially suited for performing a semiclassical approximation (expansion of the action around classical solutions) and for developing a perturbation theory with respect to some weak interaction. For $t > 0$, the path integral (2.67) obeys the Schrödinger equation. Since it is just the usual propagator (Green function), it obeys a composition law of the form

$$\langle x'', t'' | x', t' \rangle = \int_{-\infty}^{\infty} \mathrm{d}u\ \langle x'', t'' | u, \tilde{t} \rangle \langle u, \tilde{t} | x', t' \rangle. \tag{2.68}$$

This law holds because the propagator is a propagator in external time. For this reason it will not hold in quantum gravity, which is fundamentally timeless; cf. Section 5.3.4.

The quantum-mechanical path integral, which can be put on a firm mathematical footing, can be formally generalized to quantum field theory, where, however, such a footing is lacking (there is no measure-theoretical foundation). Still, the field-theoretical path integral is of great heuristic value and plays a key role, especially in gauge theories (see e.g. Böhm *et al.* (2001), among many other references). Let us consider, for example, a real scalar field $\phi(x)$.[14] Then instead of (2.66) one has the expression (setting again $\hbar = 1$)

$$Z[\phi] = \int \mathcal{D}\phi(x)\ \mathrm{e}^{\mathrm{i}S[\phi(x)]}, \tag{2.69}$$

where $Z[\phi]$ is the usual abbreviation for the path integral in the field-theoretical context (often referring to in–out transition amplitudes or to partition functions; see below). The path integral is very useful in perturbation theory and gives a concise

[14]The notation x is a shorthand for x^μ, $\mu = 0, 1, 2, 3$.

possibility for deriving Feynman rules (via the notion of the generating functional; see below). Using the methods of Grassmann integration, a path integral such as (2.69) can also be defined for fermions. However, for systems with constraints (such as gauge theories and gravity), the path-integral formulation must be generalized, as will be discussed in the course of this section. It must also be noted that the usual operator-ordering ambiguities of quantum theory are also present in the path-integral approach, in spite of the integration over classical configurations: the ambiguities are reflected here in the ambiguities in the integration measure, that is, the ambiguities related to the choice of skeletonization for the space and time intervals; see, for example, Kleinert (2009).

Instead of the original formulation in space–time, it is often appropriate to perform a rotation to four-dimensional Euclidean space via the *Wick rotation* $t \to -\mathrm{i}\tau$. In the case of the scalar field, this leads to

$$\begin{aligned} \mathrm{i}S[\phi] &= \mathrm{i}\int \mathrm{d}t\,\mathrm{d}^3x\ \left(\frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi) + \mathcal{L}_{\mathrm{int}}\right) \\ &\overset{t\to-\mathrm{i}\tau}{\longrightarrow} -\int \mathrm{d}\tau\,\mathrm{d}^3x\ \left(\frac{1}{2}\left(\frac{\partial\phi}{\partial \tau}\right)^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi) - \mathcal{L}_{\mathrm{int}}\right), \\ &=: \ -S_{\mathrm{E}}[\phi], \end{aligned} \tag{2.70}$$

where $V(\phi)$ is the potential, and an interaction $\mathcal{L}_{\mathrm{int}}$ with other fields has been taken into account. This formal rotation to Euclidean space has some advantages. First, since S_{E} is bounded from below, it improves the convergence properties of the path integral: instead of an oscillating integrand one has an exponentially damped integrand (remember, however, for example, that the Fresnel integrals used in optics are convergent in spite of the $\mathrm{e}^{\mathrm{i}x^2}$ integrand). Second, in the extremization procedure, one deals with elliptic instead of hyperbolic equations, which are more suitable for the boundary problem of specifying configurations at the initial *and* final 'times'. Third, in the Euclidean formulation the path integral can be directly related to the partition function in statistical mechanics (e.g., for the canonical ensemble, one has $Z = \mathrm{tr}\ \mathrm{e}^{-\beta H}$). Fourth, the Euclidean formulation is convenient for lattice gauge theory, in which one considers

$$Z[U] = \int \mathcal{D}U\ \mathrm{e}^{-S_{\mathrm{W}}[U]}, \tag{2.71}$$

with U denoting the lattice gauge fields defined on the links and S_{W} the Wilson action; see also Chapter 6. The justification for performing Wick rotations in quantum field theory relies on the fact that Euclidean Green functions can be analytically continued back to real time while preserving their pole structure; cf. Osterwalder and Schrader (1975).

The quantum-gravitational path integral, first formulated by Misner (1957), would be of the form

$$Z[g] = \int \mathcal{D}g_{\mu\nu}(x)\ \mathrm{e}^{\mathrm{i}S[g_{\mu\nu}(x)]}, \tag{2.72}$$

where the sum runs over all metrics on a four-dimensional manifold $\mathcal{M}$ quotiented by the diffeomorphism group $\mathrm{Diff}\mathcal{M}$. One might expect that an additional sum has to be performed over all topologies, but this is a contentious issue. As we shall see in Section 5.3.4, the path integral (2.72) behaves more like an energy Green function instead of a propagator. The reason is the absence of an external time, as already emphasized above.

Needless to say, (2.72) has a tremendously complicated nature, both technically and conceptually. One might therefore try, for the reasons stated above, to perform a Wick rotation to the Euclidean regime. This leads, however, to problems that are not present in ordinary quantum field theory. First, not every Euclidean metric (in fact, only very few) possesses a Lorentzian section, that is, leads to a signature (−,+,+,+) upon $\tau \to \mathrm{i}t$. Such a section exists only for metrics with special symmetries. (The Wick rotation is not a diffeomorphism-invariant procedure.) Second, a sum over topologies cannot be performed even in principle, because four-manifolds are not classifiable (Geroch and Hartle 1986).[15] The third, and perhaps most severe, problem is the fact that the Euclidean gravitational action is not bounded from below. Performing the same Wick rotation as above (in order to be consistent with the matter part), one finds from (1.1) the following expression for the Euclidean action:

$$S_{\mathrm{E}}[g] = -\frac{1}{16\pi G}\int_{\mathcal{M}} \mathrm{d}^4x\,\sqrt{g}\,(R-2\Lambda) - \frac{1}{8\pi G}\int_{\partial\mathcal{M}} \mathrm{d}^3x\,\sqrt{h}K. \tag{2.73}$$

To see that this action is unbounded, consider a conformal transformation of the metric, $g_{\mu\nu} \to \tilde{g}_{\mu\nu} = \Omega^2 g_{\mu\nu}$. This yields (Gibbons *et al.* 1978; Hawking 1979)

$$S_{\mathrm{E}}[\tilde{g}] = -\frac{1}{16\pi G}\int_{\mathcal{M}} \mathrm{d}^4x\,\sqrt{g}(\Omega^2 R + 6\Omega_{;\mu}\Omega_{;\nu}g^{\mu\nu} - 2\Lambda\Omega^4) - \frac{1}{8\pi G}\int_{\partial\mathcal{M}} \mathrm{d}^3x\,\sqrt{h}\Omega^2 K. \tag{2.74}$$

(Here we have transformed $K \to \mathrm{i}K$ according to the convention used in Section 4.2.1.) One recognizes that the action can be made arbitrarily negative by taking into account large gradients of the conformal factor Ω. The presence of such metrics in the path integral then leads to its divergence. This is known as the *conformal-factor problem.* There are, however, strong indications of a solution to this problem. As Dasgupta and Loll (2001) have argued, the conformal divergence can cancel against a similar term of opposite sign arising from the measure in the path integral (cf. Section 2.2.3 for a discussion of the relevant features of the measure); see Hartle and Schleich (1987) for a similar result in the context of linearized gravity. Euclidean path integrals are often used in quantum cosmology, being related to boundary conditions of the universe (see Section 8.3), so a clarification of these issues is of central interest.

Since the gravitational path integral has a highly complicated nature, the question arises whether it can be evaluated by discretization and passage to the continuum limit. In fact, the following two methods have, among others, been employed (see Section 2.2.6 for details):

1. *Regge calculus*: Originally conceived by Regge (1961) as a method for classical numerical relativity, it was applied to the Euclidean path integral from the 1980s

[15] However, it is possible that a topology change is required in quantum gravity for reasons of consistency; cf. Sorkin (1997).

on. The central idea is to decompose four-dimensional space into a set of simplices and treat the edge lengths as dynamical entities.

2. *Dynamical triangulation*: In contrast to Regge calculus, all edge lengths are kept fixed, and the sum in the path integral is instead taken over all possible manifold-gluings of equilateral simplices. The evaluation is thus reduced to a combinatorical problem. In contrast to Regge calculus, this method is applied to Lorentzian geometries, emphasizing the importance of the lightcone structure at the level of geometries in the path integral.

The discussion of path integrals will be continued in Section 2.2.3, where emphasis is put on the integration measure (the Faddeev–Popov procedure) and the derivation of Feynman rules for gravity. In the next subsection, we shall give an introduction to the use of perturbation theory in quantum gravity.

2.2.2 The perturbative non-renormalizability

In Section 2.1, we treated the concept of a graviton in the same way as the photon—within the representation theory of the Poincaré group. We have, in particular, discussed the Fierz–Pauli Lagrangian (2.20), which is (up to a total derivative) gauge invariant and which at the classical level leads inevitably to GR. The question thus arises whether this Lagrangian can be quantized in a way similar to electrodynamics, where one arrives at the very successful theory of QED. More generally, then, why should one not develop a quantum *perturbation theory* of the Einstein–Hilbert action (1.1)?

The typical situation for applications of perturbation theory in quantum field theory addresses 'scattering' situations, in which asymptotically free quantum states (representing ingoing and outgoing particles) are connected by a region of interaction. This is the standard situation in accelerators. In fact, most quantum field theories are only understood in the perturbative regime.

Perturbation theory in quantum gravity belongs to the class of covariant quantization schemes, to which also the path-integral methods belong. These methods are designed to maintain four-dimensional (space–time) covariance. They are distinguished from the canonical methods to be discussed in Chapters 4–6. Can perturbation theory be useful in quantum gravity? One might think that the gravitational interaction is intrinsically non-perturbative, and that objects such as black holes or the early universe cannot be described in perturbation theory. However, as discussed in Chapter 1, it is hopeless to probe Planck-scale effects in accelerators. While this is true, it is not excluded *per se* that perturbative quantum-gravity effects are unobservable. For example, such effects could in principle show up in the anisotropy spectrum of the cosmic microwave background; cf. Section 5.4.

A major obstacle to the viability of perturbation theory is the *non-renormalizability* of quantum GR. What does this mean? Quantum field theory uses *local* field operators $\phi(x)$. This leads to the occurrence of arbitrarily small distances and, therefore, to arbitrarily large momenta. As a consequence, *divergences* show up, usually in calculations of cross-sections coming from integrals in momentum space. A theory is said to be renormalizable if these divergences can all be removed by a redefinition of a *finite* number of physical constants (masses, charges, etc.) and fields; see e.g. Weinberg (1995) for details. These constants can only be determined experimentally.[16] A non-renormalizable theory thus needs an infinite number of parameters to be determined experimentally, which corresponds to a complete lack of predictability at the fundamental level.

It turns out that the mass dimensionality (in units for which $\hbar = c = 1$) of the coupling constant for a given interaction determines its renormalizability. This dimensionality is given by a coefficient Δ, which is called the superficial degree of divergence and which must not be negative. It can be calculated from the formula

$$\Delta := 4 - d - \sum_f n_f(s_f + 1), \tag{2.75}$$

where d is the number of derivatives; n_f is the number of fields of type f; $s_f = 0, 1/2, 1, 0$ for scalars, fermions, massive vector fields, and photons and gravitons, respectively; and Δ is equal to the mass dimension of the coupling constant. Considering, for example, the standard QED interaction $-\mathrm{i}e\bar{\psi}A^\mu\gamma_\mu\psi$, one obtains $\Delta = 4 - 0 - 3/2 - 3/2 - 1 = 0$; the electric charge e is thus dimensionless in natural units, and the coupling is renormalizable. However, the presence of a 'Pauli term' $\bar{\psi}[\gamma_\mu, \gamma_\nu]\psi F^{\mu\nu}$, for example, would lead to $\Delta = -1$ (because of the additional derivative of A^μ contained in $F^{\mu\nu}$) and thus to a non-renormalizable interaction. It was a major achievement to demonstrate that Yang–Mills theories—which are used to describe the strong and electroweak interactions—are renormalizable ('t Hooft and Veltman 1972). The Standard Model of particle physics is thus given by a renormalizable theory.

Why is the success of the Standard Model not spoiled by the presence of a non-renormalizable interaction at a large mass scale? Let us consider a non-renormalizable interaction $g \sim M^{-|\Delta|}$, where M is the corresponding mass scale. For momenta $k \ll M$, therefore, g is accompanied by a factor $k^{|\Delta|}$ to get a dimensionless number; as a consequence, this non-renormalizable interaction is suppressed by a factor $(k/M)^{|\Delta|} \ll 1$ and not seen at low momenta. The success of the renormalizable Standard Model thus indicates that any such mass scale must be much higher than currently accessible energies.

Whereas non-renormalizable theories were originally discarded as valueless, a more modern viewpoint attributes to them a possible use as *effective theories* (Weinberg 1995, Burgess 2004). If all possible terms allowed by symmetries are included in the Lagrangian, then there is a counterterm present for any ultraviolet (UV) divergence. This will become explicit in our discussion of the gravitational field below. For energies much smaller than M, effective theories might therefore lead to useful predictions. Incidentally, the Standard Model of particle physics, albeit renormalizable, is today

[16] Many physicists have considered the theory of renormalization to be only preliminary in nature, connected with a lack of understanding of the full theory: 'Some physicists may be happy to have a set of working rules leading to results in agreement with observation. They may think that this is the goal of physics. But it is not enough. One wants to understand how Nature works' (Dirac 1981), and 'Die Selbstenergie-Schwierigkeit kann nicht dadurch behoben werden, daß man ein formales Versteckenspielen mit ihr veranstaltet' (Pauli 1993).

also interpreted as an effective theory. The only truly fundamental theories are now taken to be those which unify all interactions at the Planck scale.

An early example of an effective theory is given by the Euler–Heisenberg Lagrangian,

$$\mathcal{L}_{\mathrm{E-H}} = \frac{1}{8\pi}(\mathbf{E}^2 - \mathbf{B}^2) + \frac{e^4\hbar}{360\pi^2 m_{\mathrm{e}}^4}\left[(\mathbf{E}^2 - \mathbf{B}^2)^2 + 7(\mathbf{EB})^2\right], \tag{2.76}$$

where m_{e} is the electron mass; see, for example, Section 12.3 in Weinberg (1995), or Dunne (2005). The second term in (2.76) arises after the electrons are integrated out and terms with order $\propto \hbar^2$ and higher are neglected. Already at this effective level, one can calculate observable physical effects such as Delbrück scattering (scattering of a photon by an external field).

In the *background-field method* to quantize gravity (DeWitt 1967*b*, *c*), one expands the metric about an arbitrary curved background solution to the Einstein equations,[17]

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \sqrt{32\pi G} f_{\mu\nu}. \tag{2.77}$$

Here, $\bar{g}_{\mu\nu}$ denotes the background field with respect to which (four-dimensional) covariance will be implemented in the formalism; $f_{\mu\nu}$ denotes the quantized field, which has the dimension of a mass. The covariance with respect to the background metric means that no particular background is distinguished; that is, 'background independence' is implemented in this formalism. We shall present here a heuristic discussion of the Feynman diagrams in order to demonstrate the non-renormalizability of quantum gravity; the details of the background-field method will be discussed in the following subsections.

If one chooses a flat background space–time $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$ (cf. (2.1)), one finds from the Einstein–Hilbert Lagrangian the Fierz–Pauli Lagrangian (2.20) plus higher-order terms having the symbolic form (omitting indices)

$$\sqrt{32\pi G} f(\partial f)(\partial f) + \cdots + (\sqrt{32\pi G} f)^r (\partial f)(\partial f) + \cdots \tag{2.78}$$

These are infinitely many terms because the inverse of the metric, $g^{\mu\nu}$, enters the Einstein–Hilbert Lagrangian $\propto \sqrt{-g}R_{\mu\nu}g^{\mu\nu}$. Each term contains two f-derivatives because the Ricci scalar has two derivatives; each vertex has a factor $\sqrt{G}$. Unlike the case of the Standard Model, where vertices with three and four legs occur, in this case vertices with an arbitrary number of legs are present.

Before going into the details, we shall show that the gravitational interaction is indeed non-renormalizable from the dimensional point of view. Considering the first interaction term $\sqrt{G} f(\partial f)(\partial f)$, one finds from (2.75) a negative mass dimension, $\Delta = 4 - 2 - 3(0+1) = -1$, consistent with the fact that $\sqrt{G} \propto m_{\mathrm{P}}^{-1}$.[18] Recalling the discussion after (2.49), one could say that there is a connection between the non-renormalizability of gravity and the validity of the equivalence principle.

To see this more explicitly, let us first consider the kinetic term of (2.20): after an appropriate gauge fixing (see Section 2.2.3), it leads to a propagator D of the usual form

[17]Sometimes the factor $\sqrt{8\pi G}$ or $\sqrt{16\pi G}$ is chosen instead of $\sqrt{32\pi G}$.

[18]In n space–time dimensions, $\sqrt{G}$ has dimensions of mass to the power $1 - n/2$.

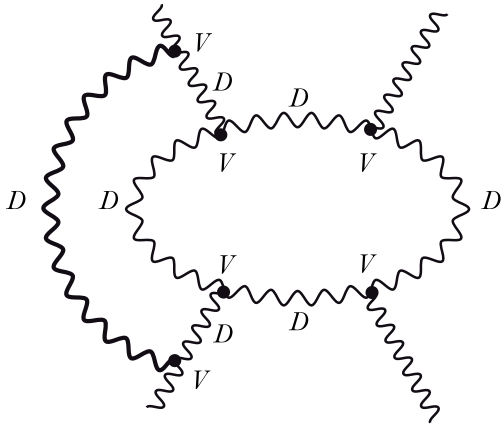

Fig. 2.2 Adding a new internal line to a Feynman diagram.

with momentum dependence $D \propto k^{-2}$. Since the interaction terms in (2.78) all contain two derivatives, one finds $V \propto k^2$ for the vertex, unlike, for example, QED, where the vertices are momentum-independent. Therefore, $DV \propto 1$ (k-independent). Consider now a Feynman diagram with one loop, r propagators, and r vertices (Fig. 2.2); cf. Duff (1975) and Deser (1989). It involves momentum integrals that together lead to the following integral (assuming for the moment n instead of four space–time dimensions):

$$\int^{p_{\mathrm{c}}} \mathrm{d}^n k \; (DV)^r \propto \int^{p_{\mathrm{c}}} \mathrm{d}^n k \propto p_{\mathrm{c}}^n ,$$

where p_{c} is some cut-off momentum. The addition of a new internal line (Fig. 2.2) then gives the new factor

$$\int^{p_{\mathrm{c}}} \mathrm{d}^n k \; D(DV)^2 \propto \int^{p_{\mathrm{c}}} \mathrm{d}^n k \; D \propto p_{\mathrm{c}}^{n-2} .$$

Therefore, if L loops are present, the degree of divergence is

$$p_{\mathrm{c}}^n \cdot p_{\mathrm{c}}^{(L-1)(n-2)} = p_{\mathrm{c}}^{(L-1)(n-2)+n} .$$

For $n = 4$, for example, this yields $p_{\mathrm{c}}^{2(L+1)}$ and thus an unbounded increase with increasing order of the diagram. In other words, since $\sqrt{G}$ has dimensions of inverse mass, $\sqrt{G} p_{\mathrm{c}}$ is dimensionless and can appear at any order. Consequently, an infinite number of divergences emerges, making the perturbation theory non-renormalizable. An exception is $n = 2$, where G is a pure number. GR is, however, trivial in two space–time dimensions, so that one can construct a sensible theory only if additional fields are added to the gravitational sector; cf. Section 5.3.5.

It must be emphasized that the counting of the degree of divergence only reflects what can be expected. This degree might well be lower due to the presence of symmetries and the ensuing cancellations of divergences. In QED, for example, the divergences are at worst logarithmic due to gauge invariance. The situation in the gravitational case will be discussed more explicitly in the next subsection.

The situation with divergences would be improved if the propagator behaved as $D \propto k^{-4}$ instead of $D \propto k^{-2}$, for then the factor corresponding to the new internal line in Fig. 2.2 would be (one also has $V \propto k^4$)

$$\int^{p_c} \mathrm{d}^n k \; D \propto p_c^{n-4}$$

and would therefore be independent of the cut-off in $n = 4$ dimensions; that is, higher loops would not lead to new divergences. This can be achieved, for example, by adding terms containing the curvature squared to the Einstein–Hilbert action because this would involve *fourth-order* derivatives. Such a theory would indeed be renormalizable, but with a high price; as Stelle (1977) has shown (and as was already noted by DeWitt (1967*b*)), the ensuing quantum theory is *not unitary*. The reason is that the propagator D can then be written in the form

$$D \propto \frac{1}{k^4 + Ak^2} = \frac{1}{A}\left(\frac{1}{k^2} - \frac{1}{k^2 + A}\right),$$

and the negative sign in front of the second term spoils unitarity: the negative sign in the propagator corresponds to a negative sign in the kinetic term, which can lead to states with negative norm and thus negative probabilities; this violates the conservation of probability. Moreover, for $A < 0$ a tachyon—a particle with a negative mass squared—can appear. For these reasons, 'exact' R^2 theories have been abandoned. However, Lagrangians with R^2 terms can, and indeed do, appear as correction terms for not-too-large curvatures (see below). (For high curvatures, R is comparable to the R^2 terms, and therefore R^3 and higher orders are also needed.) At this effective level, R^2 terms only lead to a modification of the vertices, not the propagator.

One modification of GR that leads to a perturbatively renormalizable *and* unitary quantum theory is 'Hořava–Lifshits gravity' (Hořava 2009). This is, however, achieved at the price of violating Lorentz invariance (that is, the equivalence principle) at high energies. In analogy to certain condensed-matter systems, an anisotropy is introduced between space and time. This anisotropy is described by a dynamical critical exponent z, for which the value $z = 3$ is chosen. At low energies (long distances), z assumes the value $z = 1$ and GR is regained. More details of the status of this approach can be found, for example, in Sotiriou (2011).

We now return to the perturbative quantization of GR. Since the loop expansion of Feynman diagrams is also an expansion in $\hbar$, giving $\hbar^L$ for L loops, one finds in the gravitational case a matching with the G expansion: since one recognizes from the above that for $n = 4$ there must be a factor G^{L+1} for L loops in order to compensate for the power of p_c, it is clear that $G\hbar$ is the relevant expansion parameter for pure gravity. (More precisely, one has the factor $\hbar^L G^{L-1}$ in the effective Lagrangian; see below.)

The occurrence of divergences in the quantization of gravity was first noticed by Rosenfeld (1930*a*). He calculated the gravitational energy generated by an electromagnetic field to see whether an infinite self-energy occurs in this case (the infinite self-energy for an electron had already been recognized) and found a quadratic divergence. Gupta (1952*b*) performed an expansion of the Einstein–Hilbert action around flat space–time and interpreted the non-linear terms as a direct interaction between gravitons. Like Rosenfeld, he calculated the gravitational self-energy of the photon (and also of the electron). He noted that the 'quadratically divergent term ... may be absorbed by a renormalization of the electromagnetic field strengths, but the remaining logarithmic divergent term remains a source of difficulty'. He also found hints of the non-renormalizability of quantum gravity and stated: 'This suggests that it is not possible to remove all divergencies from the gravitational field by mass and charge renormalizations'.

2.2.3 Effective action and Feynman rules

In order to apply path-integral methods to derive the effective action and Feynman diagrams, the formalism of Section 2.2.1 must be generalized to include 'gauge fixing'. There exists a general procedure, which can be found in many references (e.g. Weinberg 1996 or Böhm *et al.* 2001) and which will be briefly outlined before being applied to quantum gravity.

If we apply the general path integral in (2.72) to the expansion (2.77), we find an integral over the quantum field $f_{\mu\nu}$,

$$Z = \int \mathcal{D}f_{\mu\nu}\ \mathrm{e}^{\mathrm{i}S[f_{\mu\nu},\bar{g}_{\mu\nu}]}, \tag{2.79}$$

where in the following we shall frequently use f as an abbreviation for $f_{\mu\nu}$ (and $\bar{g}$ for the background field). The point is now that (2.79) is formally infinite because $f_{\mu\nu}$ is invariant under the gauge transformations (2.5), $f_{\mu\nu} \to f_{\mu\nu} - \partial_\nu \epsilon_\mu - \partial_\mu \epsilon_\nu =: f^{\epsilon}_{\mu\nu}$, the infinity arising from integrations over gauge-equivalent configurations. Faddeev and Popov (1967) gave a general prescription for dealing with this problem. This prescription has become especially popular in Yang–Mills theories. It consists of the following steps.

In a first step, a gauge constraint is chosen in order to fix the gauge. In the case of gravity, this would be four conditions, $G_\alpha[f,\bar{g}] = 0$. One desires to choose them such that the gauge is uniquely fixed, that is, such that each 'gauge orbit' f^ϵ is intersected exactly once. It is known, however, that this cannot always be achieved ('Gribov ambiguities'), but this problem is usually not relevant in perturbation theory. In the path integral, one then integrates over the subspace defined by $G_\alpha[f,\bar{g}] = 0$. To implement this, one defines, in a second step, a functional $\Delta_G[f,\bar{g}]$ through (neglecting in the following the dependence on the background field $\bar{g}$ for notational convenience)

$$\Delta_G[f] \cdot \int \mathcal{D}\epsilon\ \prod_\alpha \delta(G_\alpha[f^\epsilon]) = 1. \tag{2.80}$$

The integration measure is a formal integration over the gauge group and is 'left invariant', that is, $\mathcal{D}\epsilon = \mathcal{D}(\epsilon'\epsilon)$. Using this invariance of the measure, one can show that $\Delta_G[f]$ is gauge invariant, that is, $\Delta_G[f] = \Delta_G[f^\epsilon]$.

In a third step, one introduces the '1' of (2.80) into the naive path integral in (2.79). Making the substitution $f^\epsilon \to f$ and using the gauge invariance of Δ, one gets the expression

$$\int \mathcal{D}\epsilon \int \mathcal{D}f \ \prod_\alpha \delta(G_\alpha[f])\Delta_G[f]\mathrm{e}^{\mathrm{i}S[f]}.$$

The infinite term coming from the $\mathcal{D}\epsilon$ integration can now be omitted (it just corresponds to the volume of the gauge orbit). One then arrives at the following definition for the path integral (again called Z for simplicity):

$$Z := \int \mathcal{D}f \ \prod_\alpha \delta(G_\alpha[f])\Delta_G[f]\mathrm{e}^{\mathrm{i}S[f]}. \tag{2.81}$$

It depends formally on the gauge G but is in fact gauge invariant; $\Delta_G[f]$ is called the 'Faddeev–Popov determinant'.

Since a delta function appears in (2.81), we can expand $G_\alpha[f^\epsilon]$ in (2.80) around $\epsilon = 0$ to evaluate Δ_G,

$$G_\alpha[f^\epsilon] = G_\alpha[f^0] + (\hat{A}\epsilon)_\alpha,$$

where the first term on the right-hand side is zero (it is just the gauge condition), and $\hat{A}$ is the matrix (with elements $A_{\alpha\beta}$) of the derivatives of the G_α with respect to the ϵ^β. Therefore, $\Delta_G[f] = \det \hat{A}$. For the derivation of Feynman rules, it is convenient to use this expression and to rewrite the determinant as a Grassmann path integral over anticommuting fields $\eta^\alpha(x)$,

$$\det \hat{A} = \int \prod_\alpha \mathcal{D}\eta^{*\alpha}(x)\mathcal{D}\eta^\alpha(x) \ \exp\left(\mathrm{i}\int \mathrm{d}^4x \ \eta^{*\alpha}(x)A_{\alpha\beta}(x)\eta^\beta(x)\right). \tag{2.82}$$

The two independent fields $\eta^\alpha(x)$ and $\eta^{*\alpha}(x)$ are called 'vector ghosts' or 'Faddeev–Popov ghosts' because they are fermions with spin 1 and thus violate the spin–statistics theorem; they cannot appear as physical particles (external lines in a Feynman diagram) but are only introduced for mathematical convenience (and only occur inside loops in Feynman diagrams).

Besides Δ_G, the gauge-fixing part $\delta(G_\alpha)$ can also be rewritten as an effective term contributing to the action. If instead of $G_\alpha = 0$ a condition of the form $G_\alpha(x) = c_\alpha(x)$ is chosen, the corresponding path integral Z^c is in fact independent of the c_α. Therefore, if one integrates Z over c_α with an arbitrary weight function, only the (irrelevant) normalization of Z will be changed. Using a Gaussian weight function, one obtains

$$Z \propto \int \mathcal{D}c \ \exp\left(-\frac{\mathrm{i}}{4\xi}\int \mathrm{d}^4x \ c_\alpha c^\alpha\right) \int \mathcal{D}f \ \prod_\alpha \Delta_G[f]\delta(G_\alpha - c_\alpha)\mathrm{e}^{\mathrm{i}S}. \tag{2.83}$$

The c_α integration then yields

$$Z \propto \int \mathcal{D}f \ \Delta_G[f]\exp\left(\mathrm{i}S[f] - \frac{\mathrm{i}}{4\xi}\int \mathrm{d}^4x \ G_\alpha G^\alpha\right). \tag{2.84}$$

The second term in the exponential is called the 'gauge-fixing term'. Taking all contributions together, the final path integral can be written in the form (disregarding all normalization terms and reinserting the dependence on the background field)

$$Z = \int \mathcal{D}f \mathcal{D}\eta^\alpha \mathcal{D}\eta^{*\alpha} \; \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[f,\eta,\bar{g}]}, \tag{2.85}$$

where

$$\begin{aligned} S_{\mathrm{tot}}[f,\eta,\bar{g}] &= S[f,\bar{g}] - \frac{1}{4\xi}\int \mathrm{d}^4x \; G_\alpha[f,\bar{g}]G^\alpha[f,\bar{g}] \\ &\quad + \int \mathrm{d}^4x \; \eta^{*\alpha}(x) A_{\alpha\beta}[f,\bar{g}](x)\eta^\beta(x) \\ &\equiv \int \mathrm{d}^4x \; (\mathcal{L}_{\mathrm{g}} + \mathcal{L}_{\mathrm{gf}} + \mathcal{L}_{\mathrm{ghost}}). \end{aligned} \tag{2.86}$$

In this form, the path integral is suitable for the derivation of Feynman rules.

As it stands, (2.86) is also the starting point for the ‘conventional’ perturbation theory (i.e. without an expansion around a background field). In that case, the gauge-fixed Lagrangian is no longer gauge invariant but instead invariant under more general transformations including the ghost fields. This is called BRST symmetry (after the names Becchi, Rouet, Stora, and Tyutin; see e.g. DeWitt (2003) for the original references) and encodes the information about the original gauge invariance at the gauge-fixed level. We shall give a brief introduction to BRST quantization in the context of string theory later; see Chapter 9. For gravitational systems, one needs a generalization of BRST quantization known as BFV quantization; see Batalin and Vilkovisky (1977) and Batalin and Fradkin (1983). This generalization is, for example, responsible for the occurrence of a four-ghost vertex in perturbation theory.

In the background-field method, on the other hand, the gauge symmetry (here, symmetry with respect to coordinate transformations) is preserved for the background field $\bar{g}$. For this reason, one can call this method a *covariant* approach to quantum gravity. The formal difference between the ‘conventional’ and the background-field method is in the form of the gauge-fixing terms (Duff 1975, Böhm *et al.* 2001), leading to different diagrams in this sector. It turns out that, once the Feynman rules are obtained, calculations are often simpler in the background-field method. For non-gauge theories, the two methods are identical. The background-field and BRST methods are alternative procedures to arrive at the same physical results.

It is, of course, a non-trivial task to investigate whether path-integral quantization is equivalent to other methods of quantization. In fact, such an equivalence does not hold in general; see in this context the instructive example discussed by Klauder (2001, p. 147).

Let us now summarize the application of these methods to the quantization of the gravitational field. For vanishing cosmological constant, the Lagrangian is given in this case by (see ’t Hooft and Veltman 1974)

$$\mathcal{L}_{\mathrm{g}} = \frac{\sqrt{-g}R}{16\pi G} = \sqrt{-\bar{g}}\left(\frac{\bar{R}}{16\pi G} + L_{\mathrm{g}}^{(1)} + L_{\mathrm{g}}^{(2)} + \dots\right), \tag{2.87}$$

where the ‘barred’ quantities refer to the background metric; see (2.77). For $L_{\mathrm{g}}^{(1)}$, one has the expression

$$L_{\rm g}^{(1)} = \frac{f_{\mu\nu}}{\sqrt{32\pi G}} \left(\bar{g}^{\mu\nu}\bar{R} - 2\bar{R}^{\mu\nu}\right). \tag{2.88}$$

This vanishes if the background is a solution of the (vacuum) Einstein equations. The expression for $L_{\rm g}^{(2)}$ reads

$$\begin{aligned} L_{\rm g}^{(2)} = {} & \tfrac{1}{2}\, f_{\mu\nu;\alpha} f^{\mu\nu;\alpha} - \tfrac{1}{2}\, f_{;\alpha} f^{;\alpha} + f_{;\alpha} f^{\alpha\beta}{}_{;\beta} - f_{\mu\beta;\alpha} f^{\mu\alpha;\beta} \\ & + \bar{R}(\tfrac{1}{2}\, f_{\mu\nu} f^{\mu\nu} - \tfrac{1}{4}\, f^2) + \bar{R}^{\mu\nu}\left(f f_{\mu\nu} - 2 f_\mu{}^\alpha f_{\nu\alpha}\right). \end{aligned} \tag{2.89}$$

The first line corresponds to the Fierz–Pauli Lagrangian (2.20), while the second line describes the interaction with the background (not present in (2.20) because there the background is flat); we recall that $f := f_\mu{}^\mu$. For the gauge-fixing part $\mathcal{L}_{\rm gf}$, one chooses

$$\mathcal{L}_{\rm gf} = \sqrt{-\bar{g}}(f_{\mu\nu}{}^{;\nu} - \tfrac{1}{2}\, f_{;\mu})(f^{\mu\rho}{}_{;\rho} - \tfrac{1}{2}\, f^{;\mu}). \tag{2.90}$$

This condition corresponds to the 'harmonic gauge condition' (2.3) and turns out to be a convenient choice because the terms containing the second derivatives of $f_{\mu\nu}$ take the form of a d'Alembertian. For the ghost part, one finds

$$\mathcal{L}_{\rm ghost} = \sqrt{-\bar{g}}\ \eta^{*\mu}\left(\eta_{\mu;\sigma;}{}^{\sigma} - \bar{R}_{\mu\nu}\eta^\nu\right). \tag{2.91}$$

The first term in the brackets is the covariant d'Alembertian, $\Box\eta_\mu$. The full action in the path integral (2.85) then reads, with the background metric obeying Einstein's equations,

$$\begin{aligned} S_{\rm tot} = \int {\rm d}^4x\, \sqrt{-\bar{g}} \Bigg(& \frac{\bar{R}}{16\pi G} - \frac{1}{2} f_{\mu\nu} D^{\mu\nu\alpha\beta} f_{\alpha\beta} + \frac{f_{\mu\nu}}{\sqrt{32\pi G}}\left[\bar{g}^{\mu\nu}\bar{R} - 2\bar{R}^{\mu\nu}\right] \\ & + \eta^{*\mu}(\bar{g}_{\mu\nu}\Box - \bar{R}_{\mu\nu})\eta^\nu + \mathcal{O}(f^3)\Bigg), \end{aligned} \tag{2.92}$$

where $D^{\mu\nu\alpha\beta}$ is a shorthand for the terms occurring in (2.89) and (2.90). The desired Feynman diagrams can then be obtained from this action. In contrast to the original action without gauge fixing, the operator $D^{\mu\nu\alpha\beta}$ is invertible and defines both propagator and vertices (interaction with the background field). The explicit expressions are complicated; see 't Hooft and Veltman (1974) and Donoghue (1994). They simplify considerably in the case of a flat background, for which, for example, the propagator (in the harmonic gauge) is given by the expression

$$D_{\mu\nu\alpha\beta} = \frac{1}{2(k^2 - {\rm i}\epsilon)}(\eta_{\mu\alpha}\eta_{\nu\beta} + \eta_{\mu\beta}\eta_{\nu\alpha} - \eta_{\mu\nu}\eta_{\alpha\beta}). \tag{2.93}$$

(The structure of the term in brackets is reminiscent of a four-dimensional version of the DeWitt metric discussed in Section 4.1.2.) The action (2.92) leads to diagrams with at most one loop such as those depicted in Fig. 2.3 (the 'one-loop approximation'). Figure 2.3(a) describes a graviton loop interacting with the background field. It is a virtue of the background-field method that an arbitrary number of external lines can be considered. Figure 2.3(b) describes a ghost loop interacting with the background

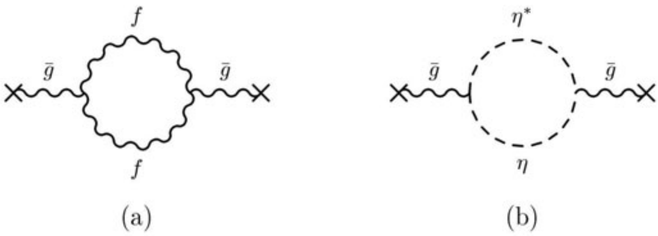

Fig. 2.3 (a) Graviton loop interacting with the background field. (b) Ghost loop interacting with the background field.

field. Ghosts are needed to guarantee the unitarity of the S-matrix, as was already noted by Feynman (1963) and DeWitt (1967*b*).

We note that in n space–time dimensions the propagator (2.93) reads

$$D_{\mu\nu\alpha\beta} = \frac{1}{2(k^2 - \mathrm{i}\epsilon)}(\eta_{\mu\alpha}\eta_{\nu\beta} + \eta_{\mu\beta}\eta_{\nu\alpha} - \frac{2}{n-2}\eta_{\mu\nu}\eta_{\alpha\beta}).$$

The pole for $n = 2$ arises because in two space–time dimensions, GR is a topological theory.

The use of the background-field method guarantees covariance with respect to the background field. This covariance is achieved in particular by the implementation of dimensional regularization to treat the divergences that arise.[19] The divergences are local and give rise to covariant terms (such as R, R^2, etc.), which must be absorbed by terms of the same form already present in the Lagrangian. Since curvature terms different from R are absent in the original Einstein–Hilbert Lagrangian,[20] one must start not with $\mathcal{L}_{\mathrm{g}}$ but with a Lagrangian of the form

$$\tilde{\mathcal{L}}_{\mathrm{g}} = \mathcal{L}_{\mathrm{g}} + \mathcal{L}_{(2)} + \mathcal{L}_{(3)} + \ldots + \mathcal{L}_{\mathrm{m}}, \tag{2.94}$$

where

$$\begin{aligned} \mathcal{L}_{(2)} &= \sqrt{-\bar{g}}(c_1\bar{R}^2 + c_2\bar{R}_{\mu\nu}\bar{R}^{\mu\nu}), \\ \mathcal{L}_{(3)} &= \mathcal{O}(\bar{R}^3)\ , \ \text{etc.}, \end{aligned} \tag{2.95}$$

and $\mathcal{L}_{\mathrm{m}}$ denotes the Lagrangian for non-gravitational fields. Note that because of the Gauss–Bonnet theorem, the expression

$$\int \mathrm{d}^4x\ \sqrt{-g}(R_{\mu\nu\sigma\tau}R^{\mu\nu\sigma\tau} - 4R_{\mu\nu}R^{\mu\nu} + R^2)$$

is a topological invariant, which is why a term proportional to $R_{\mu\nu\sigma\tau}R^{\mu\nu\sigma\tau}$ does not have to be considered.

[19] Dimensional regularization employs an auxiliary extension of the number n of space–time dimensions in a continuous way such that no divergences appear for $n \neq 4$. The divergences that occur for $n \to 4$ are then subtracted.

[20] This is different from QED, where the only terms that occur have the same form as terms already present in the Maxwell Lagrangian. The same holds for Yang–Mills theories.

Can one justify the use of the modified gravitational Lagrangian (2.94)? Higher powers of the curvature emerge generically from fundamental theories in the form of effective theories.[21] This is the case, for example, in string theory (Chapter 9). Thus it has to be expected that the Einstein–Hilbert action (1.1) is not exact even at the classical level, but must be augmented by an action containing higher-order curvature terms. Experimental limits on the parameters c_1, c_2, etc., are very weak because curvatures are usually very small (for example, in the solar system). Stelle (1978) has estimated that from the perihelion motion of Mercury one gets only $c_1, c_2 \lesssim 10^{88}$. As one would expect c_1 and c_2 to be much less than that, there is no hope of measuring them in solar system experiments.

With (2.92) substituted in the path integral, integration over the $f_{\mu\nu}$ gives a term proportional to $(\det D)^{-1/2} = \exp(-\frac{1}{2}\mathrm{tr}\ln D)$, where D is a shorthand notation for $D^{\mu\nu\alpha\beta}$ ('t Hooft and Veltman 1974). In the one-loop approximation used here, this yields a divergent contribution to the Lagrangian of the form

$$\mathcal{L}^{(\mathrm{div})}_{1-\mathrm{loop}} = \frac{\hbar}{8\pi^2\epsilon}\sqrt{-\bar{g}}\left(\frac{\bar{R}^2}{120} + \frac{7}{20}\bar{R}_{\mu\nu}\bar{R}^{\mu\nu}\right), \tag{2.96}$$

where $\epsilon = 4 - n$ is the parameter occurring in dimensional regularization; (2.96) diverges in the limit of space–time dimension $n \to 4$. For pure gravity, $\mathcal{L}^{(\mathrm{div})}_{1-\mathrm{loop}}$ is zero if the background field is 'on-shell', that is, a solution of the (vacuum) Einstein equations. This is a feature of the background-field method. In the presence of matter, this no longer holds. The one-loop counterterms (2.96) can be eliminated by a local redefinition of the graviton field $f_{\mu\nu}$ ('t Hooft and Veltman 1974, Krasnov 2010).

Adding (2.96) to (2.94), one obtains the 'renormalized' values for c_1 and c_2,

$$c_1^{(\mathrm{ren})} = c_1 + \frac{\hbar}{960\pi^2\epsilon}\ , \quad c_2^{(\mathrm{ren})} = c_2 + \frac{7\hbar}{160\pi^2\epsilon}. \tag{2.97}$$

These constants can only be 'measured' (or determined from a fundamental theory); here their use is just to absorb the divergences present in (2.96).

If only a non-self-interacting field is present in addition to the gravitational field, the one-loop approximation for this field is exact. It is important to emphasize that in this approximation the effect of the gravitons must be taken into account on an equal footing with matter, since no G appears in the one-loop Lagrangian; see (2.96). This is why the terms 'quantum field theory in curved space–time' and 'one-loop quantum gravity' are both used for this level of approximation.

The divergences at two-loop order were first calculated by Goroff and Sagnotti (1985) using computer methods. Their result was later confirmed by van de Ven (1992). An analytic calculation of two-loop divergences was performed by Barvinsky and Vilkovisky (1987). The divergent part in the two-loop Lagrangian was found to be

$$\mathcal{L}^{(\mathrm{div})}_{2-\mathrm{loop}} = \frac{209\hbar^2}{2880}\frac{32\pi G}{(16\pi^2)^2\epsilon}\sqrt{-\bar{g}}\bar{R}^{\alpha\beta}{}_{\gamma\delta}\bar{R}^{\gamma\delta}{}_{\mu\nu}\bar{R}^{\mu\nu}{}_{\alpha\beta}. \tag{2.98}$$

The divergence as $\epsilon \to 0$ can be absorbed by a corresponding term in $\mathcal{L}_{(3)}$ of (2.95). In contrast to (2.96), the gravitational constant G occurs here. The two-loop divergence

[21] To quote from Weinberg (1997), 'Why would anyone suppose that these higher terms are absent?'

can no longer be removed by a local redefinition of the graviton field, but one can show that it can be removed by a *non-local* field redefinition (Krasnov 2010).

The two-loop divergence (2.98) is also present for pure gravity, because it contains the full Riemann tensor and not just the Ricci tensor and Ricci scalar as the one-loop divergence (2.96) does. Since the Riemann tensor vanishes for Minkowski space, one might think that no divergences arise at, for example, the two-loop level for graviton scattering in Minkowski space. But this is not what (2.98) implies. When one calculates the scattering of particles on some background, one has to substitute into the effective action not the pure background solution, but the field excitations over the background as induced by the in- and out-states that enter into the matrix element to be calculated. While the Ricci tensor vanishes for these excitations, the Riemann tensor does not. Thus the divergences for graviton scattering on Minkowski space vanish at the one-loop, but not at the two-loop level.

For higher loops, one finds from dimensional analysis that the divergent part of the Lagrangian is of the form

$$\mathcal{L}^{(\mathrm{div})}_{L-\mathrm{loop}} \sim \sqrt{-\bar{g}}\hbar^L G^{L-1} \nabla^p \bar{R}^m \frac{1}{\epsilon}\ , \quad p+m=L+1 \tag{2.99}$$

(recall that the loop expansion is also a WKB expansion), where the symbolic notation $\nabla^p \bar{R}^m$ is a shorthand for all curvature terms and their derivatives that can occur at this order.

Divergences remain if other fields (scalars, photons, Yang–Mills fields) are coupled (Deser *et al.* 1975). In the Einstein–Maxwell theory, for example, one obtains instead of (2.96) the expression

$$\mathcal{L}^{(\mathrm{div})}_{1-\mathrm{loop}} = \sqrt{-\bar{g}}\frac{137\hbar}{60\epsilon}\bar{R}_{\mu\nu}\bar{R}^{\mu\nu}. \tag{2.100}$$

We shall see in Section 2.3 how supergravity can improve the situation without, however, avoiding the occurrence of divergences.

The above treatment of divergences concerns the (ill-understood) UV behaviour of the theory and does not lead to any new predictions.[22] Genuine predictions can, however, be obtained from the action (2.92) in the infrared limit (Donoghue 1994). These predictions are independent of the unknown coefficients c_1, c_2, etc. Let us consider two examples. The first example is a quantum-gravitational correction term to the Newtonian potential[23]

$$V(r) = -\frac{Gm_1m_2}{r}$$

between two masses m_1 and m_2. It is appropriate to define the potential through the scattering amplitude. The non-analytic parts of this amplitude give the long-range, low-energy corrections to the Newtonian potential. 'Non-analytic' here refers to terms

[22] It has, however, been suggested that in certain situations theories with infinitely many couplings can be studied perturbatively also at high energies; cf. Anselmi (2003).

[23] The Newtonian potential was derived from linear quantum gravity by Bronstein (1936).

such as $\ln(-p^2)$ and $(-p^2)^{-1/2}$, where p^2 denotes the square of the four-momentum. After a rather long calculation, Bjerrum-Bohr *et al.* (2003*a*) found (restoring c)

$$V(r) = -\frac{Gm_1m_2}{r}\left(1 + 3\frac{G(m_1+m_2)}{rc^2} + \frac{41}{10\pi}\frac{G\hbar}{r^2c^3} + \mathcal{O}(G^2)\right). \tag{2.101}$$

All terms are fully determined by the non-analytic parts of the one-loop amplitude; the parameters connected with the higher curvature terms in the action contribute only to the analytic parts. It is for this reason that an unambiguous result can be obtained. Note that (2.101) corresponds to an effective gravitational constant $G_{\text{eff}}(r) > G$. Results such as (2.101) can, of course, only be trusted if effects from two and higher loops do not spoil them (which is presently unknown).

Although arising from a one-loop amplitude, the first correction term is in fact an effect of classical GR. It can be obtained from the Einstein–Infeld–Hoffmann equations, in which neither of the two bodies is treated as a test body. Interestingly, such a term had already been derived from quantum-gravitational considerations by Iwasaki (1971).

The second correction is proportional to $\hbar$ and is of genuine quantum-gravitational origin. The sign in front of this term indicates that the strength of the gravitational interaction is increased as compared to the pure Newtonian potential.[24] The result (2.101) demonstrates that a definite prediction from quantum gravity is in principle possible. Unfortunately, the correction term, being of the order $(l_{\text{P}}/r)^2 \ll 1$, is not measurable in laboratory experiments: taking the Bohr radius for r, the correction is of order 10^{-49}. We remark that similar techniques have been applied successfully in low-energy QCD (in the limit of pion masses $m_\pi \to 0$) and are known under the term 'chiral perturbation theory' (which is also a 'non-renormalizable theory' with a dimensionful coupling constant); see, for example, Gasser and Leutwyler (1984). Quantum corrections to the Schwarzschild and Kerr metrics were calculated along these lines in Bjerrum-Bohr *et al.* (2003*b*).

One can also calculate by these methods the quantum-gravitational corrections terms to the Coulomb potential (Faller 2008). The dominating terms turn out to be

$$V(r) = \frac{Q_1Q_2}{r}\left(1 + 3\frac{G(m_1+m_2)}{rc^2} + \frac{6}{\pi}\frac{G\hbar}{r^2c^3}\right) + \dots\ ,$$

with the neglected terms containing in addition a factor $\alpha = e^2/\hbar c$.

The second example is graviton–graviton scattering. This is the simplest low-energy process in quantum gravity. It was originally calculated at tree level by a student of Bryce DeWitt; see DeWitt (1967*c*) and DeWitt-Morette (2011). For the scattering of a graviton with helicity $+2$ on a graviton with helicity -2, for example, he found for the cross-section in the centre-of-mass frame the expression

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = 4G^2E^2\frac{\cos^{12}\theta/2}{\sin^4\theta/2}, \tag{2.102}$$

[24]There had been some disagreement about the exact number in (2.101); see the discussion in Bjerrum-Bohr *et al.* (2003*a*).

where E is the centre-of-mass energy (and similar results for other combinations of helicity). One recognizes in the denominator of (2.102) the term well known from Rutherford scattering. For this calculation, more than 500 terms arising from four Feynman diagrams had to be calculated. The same result can be obtained from one diagram in string theory involving closed strings (Sannan 1986); see also Section 9.2. DeWitt (1967*c*) also discussed other processes such as gravitational bremsstrahlung.

One-loop calculations can also be carried out. In the background-field method, the quantum fields $f_{\mu\nu}$ occur only in internal lines; external lines contain only the background field $\bar{g}_{\mu\nu}$. It has already been mentioned that this makes the whole formalism 'covariant'. Donoghue and Torma (1999) have shown that one-loop calculations of graviton–graviton scattering yield a finite result in the infrared (IR) limit; the result is independent of any parameters such as c_1 or c_2. The cancellation of IR divergences due to the emission of soft gravitons is needed and has been shown in fact to occur (as, for example, in QED). This yields again a definite result from quantum gravity. There is, in fact, a huge literature about IR effects. One example is the dynamical relaxation of the cosmological constant and its possible relevance to the dark-matter problem (Tsamis and Woodard 1993). Another way of addressing this issue is through the investigation of renormalization-group equations, which can be applied also to effective theories; cf. Section 2.2.5. At least in principle, this method could explain the occurrence of a small positive cosmological constant in agreement with observations, because it would arise as a strong IR quantum effect.

The idea that non-renormalizable theories can be treated as ordinary physical theories from which phenomenological consequences can be drawn has also been discussed in other contexts. Kazakov (1988), for example, generalized the standard formalism of the renormalization group to non-renormalizable theories. Barvinsky *et al.* (1993) developed a version of the renormalization-group formalism for non-renormalizable theories, which is particularly convenient for applications to GR coupled to a scalar field.

The inclusion of gravity can also dramatically change the renormalization-group equation for the electric charge. Toms (2010) has found that the electric charge vanishes for high energies if the coupling to gravity is taken into account. More precisely, for the energy-dependent fine-structure constant $\alpha(E)$, he derived the renormalization-group equation (see also (2.130) below)

$$E\frac{\mathrm{d}\alpha(E)}{\mathrm{d}E} = \frac{\alpha^2}{6\pi^2} - \frac{2G\alpha}{\pi^2\hbar}\left(E^2 + \frac{3\hbar^2\Lambda}{2}\right),$$

where Λ is the cosmological constant. Without gravity, the second term on the right-hand side would be absent, and one would recover the standard result of the electric charge increasing with energy. The inclusion of gravity leads to the minus sign in front of the second term and thus leads to a decreasing electric charge for high energies—QED then becomes asymptotically free. This example demonstrates what far-reaching consequences the influence of gravity could have even on standard theories such as QED.

2.2.4 Semiclassical Einstein equations

In this subsection, we shall give a general introduction to the concept of the effective action, which is of central importance for quantum field theory. We shall then apply this to quantum gravity and present in particular a derivation of the semiclassical Einstein equations (1.37). More details can be found in Barvinsky (1990) and Buchbinder *et al.* (1992).[25]

For a general quantum field φ (with possible components φ^i), the *generating functional* $W[J]$ is defined by the path integral

$$\langle \mathrm{out}, 0|\mathrm{in}, 0\rangle_J =: Z[J] =: \mathrm{e}^{\mathrm{i}W[J]} = \int \mathcal{D}\varphi\ \mathrm{e}^{\mathrm{i}S[\varphi]+\mathrm{i}J_k\varphi^k}, \tag{2.103}$$

where J is an external current and $J_k\varphi^k$ is an abbreviation for $\int \mathrm{d}^4x J_i(x)\varphi^i(x)$. This is known as 'DeWitt's condensed notation'; see DeWitt (1965). We usually write the index of φ only in expressions where more than one index occurs.

If φ is a gauge field, the measure in (2.103) is understood as including gauge-fixing terms and Faddeev–Popov ghosts; see (2.86) above. Later, φ will be the gravitational field. $W[J]$ is called the generating functional because one can calculate Green functions of the theory from it. More precisely, W generates *connected* Green functions[26] according to

$$\langle \varphi_1 \cdots \varphi_k\rangle_J = \mathrm{e}^{-\mathrm{i}W[J]}\left(\frac{1}{\mathrm{i}}\right)^k \frac{\delta^k}{\delta J_1\cdots\delta J_k}\mathrm{e}^{\mathrm{i}W[J]}, \tag{2.104}$$

where the expectation value occurring on the left-hand side is defined by the expression

$$\langle A(\varphi)\rangle := \frac{\langle \mathrm{out}, 0|T(A(\varphi))|\mathrm{in}, 0\rangle}{\langle \mathrm{out}, 0|\mathrm{in}, 0\rangle}, \tag{2.105}$$

where T is time ordering.[27] The expectation value of the field in the presence of the external current (the *mean field*) is given by the expression

$$\langle\varphi\rangle_J = Z^{-1}\int \mathcal{D}\varphi\ \varphi\ \mathrm{e}^{\mathrm{i}S[\varphi]+\mathrm{i}J_k\varphi^k}.$$

One can then write

$$\langle\varphi\rangle_J = \frac{\delta W[J]}{\delta J}. \tag{2.106}$$

The two-point function then follows from (2.104),

$$\begin{aligned}\langle\varphi_i\varphi_k\rangle_J &= -\mathrm{i}\frac{\delta^2 W}{\delta J_i \delta J_k} + \langle\varphi_i\rangle_J\langle\varphi_k\rangle_J \\ &=: -\mathrm{i}G_J^{ik} + \langle\varphi_i\rangle_J\langle\varphi_k\rangle_J.\end{aligned} \tag{2.107}$$

[25] In the preparation of this subsection, I have benefited much from discussions with Andrei Barvinsky.

[26] A connected Green function is a Green function referring to a connected graph, that is, a graph for which any two of its points are connected by internal lines.

[27] For fermionic fields, one must distinguish between functional derivatives from the left and from the right.

The 'propagator' that yields the complete two-point function of the theory is then given by G_J^{ik} evaluated at $J = 0$ and is simply denoted by G^{ik}.

The central concept in this formalism is the *effective action*, $\Gamma[\langle\varphi\rangle]$, which is a functional of the mean field. It is defined here by the Legendre transformation

$$\Gamma[\langle\varphi\rangle] = W[J] - \int \mathrm{d}^4x \; J(x)\langle\varphi(x)\rangle, \tag{2.108}$$

where J is expressed through $\langle\varphi\rangle$ (inversion of (2.106)). It therefore follows that

$$\frac{\delta\Gamma}{\delta\langle\varphi\rangle} = -J(\langle\varphi\rangle). \tag{2.109}$$

In the absence of external sources this is

$$\frac{\delta\Gamma}{\delta\langle\varphi\rangle} = 0, \tag{2.110}$$

generalizing the classical equations $\delta S/\delta\varphi_{\mathrm{cl}} = 0$. Equation (2.110) describes the dynamics of the mean field *including all quantum corrections*, since in the absence of external sources one has

$$\mathrm{e}^{\mathrm{i}\Gamma[\langle\varphi\rangle]} = \int \mathcal{D}\varphi \; \mathrm{e}^{\mathrm{i}S[\varphi]}.$$

That Γ contains all information about the full theory can also be understood as follows: solve (2.109) for $\langle\varphi\rangle$ and insert it into (2.107). This yields the two-point function, and a similar procedure leads to all higher correlation functions. Unfortunately, Γ is not invariant under field redefinitions, although the S-matrix is; see Barvinsky (1990).

One can show that Γ generates all one-particle irreducible Feynman diagrams.[28] The irreducible part of the two-point function is given by

$$D_{ik} = \frac{\delta^2\Gamma}{\delta\langle\varphi^i\rangle\delta\langle\varphi^k\rangle} = -\frac{\delta J(\langle\varphi_k\rangle)}{\delta\langle\varphi^i\rangle}. \tag{2.111}$$

From (2.107), one has $G_J^{ik} D_{kl} = -\delta_l^i$.

The effective action is the appropriate generalization of the classical action to quantum theory. This is exhibited in particular if one performs a semiclassical ('loop') expansion. This will now be discussed in some detail. For this purpose, we reintroduce $\hbar$ into the formalism. Multiplication of (2.103) on both sides by (again using DeWitt's condensed notation)

$$\exp\left(-\mathrm{i}J\langle\varphi\rangle/\hbar\right) = \exp\left(\mathrm{i}\frac{\delta\Gamma}{\delta\langle\varphi\rangle}\langle\varphi\rangle/\hbar\right)$$

gives

$$\exp\left(\frac{\mathrm{i}}{\hbar}\Gamma[\langle\varphi\rangle]\right) = \int \mathcal{D}\varphi \; \exp\left(\frac{\mathrm{i}}{\hbar}S[\varphi] - \frac{\mathrm{i}}{\hbar}\frac{\delta\Gamma}{\delta\langle\varphi\rangle}(\varphi - \langle\varphi\rangle)\right). \tag{2.112}$$

This is an *exact* equation for the effective action. It will be solved iteratively by an expansion with respect to $\hbar$ ('stationary-phase approximation'). Expanding the classical

[28] A diagram is called irreducible if it does not decompose into two separate diagrams when just one internal line is cut.

action around the mean field yields (recall that we write the index for φ usually only when at least two different indices occur)

$$S[\varphi] = S[\langle\varphi\rangle] + \frac{\delta S}{\delta\langle\varphi\rangle}(\varphi - \langle\varphi\rangle) + \frac{1}{2}\frac{\delta^2 S}{\delta\langle\varphi_i\rangle\delta\langle\varphi_k\rangle}(\varphi_i - \langle\varphi_i\rangle)(\varphi_k - \langle\varphi_k\rangle) + \dots .$$

Writing

$$\Gamma[\langle\varphi\rangle] = S[\langle\varphi\rangle] + \Gamma_{\text{loop}}[\langle\varphi\rangle]$$

and introducing $\Delta_k := \varphi_k - \langle\varphi_k\rangle$, one gets from (2.112) the expression

$$\exp\left(\frac{\mathrm{i}}{\hbar}\Gamma_{\text{loop}}[\langle\varphi\rangle]\right) = \int \mathcal{D}\Delta \exp\left(\frac{\mathrm{i}}{2\hbar}\frac{\delta^2 S}{\delta\langle\varphi_i\rangle\delta\langle\varphi_k\rangle}\Delta_i\Delta_k\right) \times \left(1 - \frac{\mathrm{i}}{\hbar}\frac{\delta\Gamma_{\text{loop}}}{\delta\langle\varphi\rangle}\Delta + \frac{\mathrm{i}}{3!\hbar}S^{(3)}\Delta\Delta\Delta + \frac{\mathrm{i}}{4!\hbar}S^{(4)}\Delta\Delta\Delta\Delta + \left(\frac{\mathrm{i}}{3!\hbar}\right)^2 [S^{(3)}\Delta\Delta\Delta]^2 + \dots\right). \tag{2.113}$$

We have introduced here the abbreviation

$$S^{(3)}\Delta\Delta\Delta := \frac{\delta^3 S}{\delta\langle\varphi_i\rangle\delta\langle\varphi_k\rangle\delta\langle\varphi_l\rangle}\Delta_i\Delta_k\Delta_l$$

(which also includes an integration), and similar abbreviations for the other terms. We emphasize that the first functional derivative of S with respect to the mean field has cancelled in (2.113); this derivative is not zero (as sometimes claimed) because the mean field is at this stage arbitrary and does not have to satisfy the classical field equations. In the integral over Δ, the terms odd in Δ vanish. Moreover, after this integration the terms of fourth and sixth power in Δ yield terms proportional to $\hbar^2$ and $\hbar^3$, respectively.

In general, one has the loop expansion

$$\Gamma_{\text{loop}} = \sum_{L=1}^{\infty} \Gamma^{(L)}[\langle\varphi\rangle], \tag{2.114}$$

where $\Gamma^{(L)}$ is of order $\hbar^L$. In the highest ('one loop') order for $\Gamma^{(1)}$ one thus has to evaluate only the integral over the exponential in (2.113). This yields

$$\Gamma^{(1)} = \frac{\mathrm{i}\hbar}{2}\ln\det\frac{\delta^2 S[\langle\varphi\rangle]}{\delta\langle\varphi\rangle\delta\langle\varphi\rangle} + \mathcal{O}(\hbar^2). \tag{2.115}$$

We now want to investigate (2.109) at one-loop order. Introducing the notation

$$\frac{\delta^2 S[\langle\varphi\rangle]}{\delta\langle\varphi_i\rangle\delta\langle\varphi_k\rangle} =: S_{ik},$$

we get

$$\begin{aligned}\frac{\delta\Gamma^{(1)}}{\delta\langle\varphi_j\rangle} &= \frac{\mathrm{i}\hbar}{2}\left(\det S_{ik}\right)^{-1}\frac{\delta(\det S_{mn})}{\delta\langle\varphi_j\rangle} \\ &= \frac{\mathrm{i}\hbar}{2}\left(S^{-1}\right)^{mn}\frac{\delta S_{nm}}{\delta\langle\varphi_j\rangle} \equiv -\frac{\mathrm{i}\hbar}{2}\mathcal{G}^{mn}\frac{\delta S_{nm}}{\delta\langle\varphi_j\rangle},\end{aligned} \tag{2.116}$$

where $\mathcal{G}^{mn}$ denotes the propagator occurring in (2.107) evaluated at this order of approximation, and the identity

$$\frac{\delta(\det A)}{\det A} = \mathrm{tr}\left(A^{-1}\delta A\right)$$

has been used at the beginning of the second line. We introduce the notation

$$\frac{\delta S_{nm}}{\delta\langle\varphi_j\rangle} =: S_{jmn}\ , \quad S_{ik} =: F(\nabla)\delta(x,y),$$

where we have for simplicity suppressed possible discrete indices attached to the differential operator F. (For example, for a free massless scalar field we just have $F = -\Box$.) This may look confusing, but in DeWitt's condensed notation discrete and continuous indices appear on the same footing. Then,

$$\mathcal{G}^{mn} \equiv \mathcal{G}(x,y)\ , \quad F(\nabla)\mathcal{G}(x,y) = -\delta(x,y),$$

and (2.116) reads

$$\frac{\delta\Gamma^{(1)}}{\delta\langle\varphi(x)\rangle} = -\frac{\mathrm{i}\hbar}{2}\int \mathrm{d}y\,\mathrm{d}z\ \mathcal{G}(y,z)\frac{\delta^3 S}{\delta\langle\varphi(z)\rangle\delta\langle\varphi(y)\rangle\delta\langle\varphi(x)\rangle} \equiv -\frac{\mathrm{i}\hbar}{2}\mathcal{G}^{mn}S_{nmk}. \tag{2.117}$$

From (2.109), we then get the effective field equation up to one-loop order,

$$\frac{\delta S}{\delta\langle\varphi(x)\rangle} - \frac{\mathrm{i}\hbar}{2}\mathcal{G}^{mn}S_{nmk} = -J(x). \tag{2.118}$$

The second term on the left-hand side thus yields the *first quantum correction* to the classical field equations. According to the Feynman rules, it corresponds to a one-loop diagram.

So far, the formalism applies to any quantum field. Let us now switch to quantum gravity, where $\varphi(x)$ corresponds to $(g_{\mu\nu}(x),\ \phi(x))$, in which $\phi(x)$ represents a non-gravitational field. For the classical part in (2.118), we get, writing $S = S_{\mathrm{EH}} + S_{\mathrm{m}}$,

$$\frac{\delta S_{\mathrm{EH}}}{\delta\langle g^{\mu\nu}(x)\rangle} + \frac{\delta S_{\mathrm{m}}}{\delta\langle g^{\mu\nu}(x)\rangle} = \frac{\sqrt{-g}}{16\pi G}\left((R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) - \frac{\sqrt{-g}}{2}T_{\mu\nu}, \tag{2.119}$$

where the right-hand side is evaluated at the mean metric $\langle g_{\mu\nu}\rangle$. For the quantum correction in (2.118), one gets the sum of a 'matter loop' and a 'graviton loop',

$$
\begin{aligned}
-\frac{\mathrm{i}\hbar}{2}\mathcal{G}^{mn}S_{nmk} \equiv & -\frac{\mathrm{i}\hbar}{2}\int \mathrm{d}z\,\mathrm{d}y\ \frac{\delta^3 S}{\delta\langle g^{\mu\nu}(x)\rangle\delta\langle\phi(z)\rangle\delta\langle\phi(y)\rangle}\mathcal{G}_{\mathrm{m}}(z,y) \\
& -\frac{\mathrm{i}\hbar}{2}\int \mathrm{d}z\,\mathrm{d}y\ \frac{\delta^3 S}{\delta\langle g^{\mu\nu}(x)\rangle\delta\langle g^{\alpha\beta}(z)\rangle\delta\langle g^{\gamma\delta}(y)\rangle}\mathcal{G}^{\alpha\beta,\gamma\delta}(z,y). \qquad (2.120)
\end{aligned}
$$

Using again the expression for the energy–momentum tensor as a variational derivative of the matter action, one recognizes that the matter loop is given by

$$
\frac{\mathrm{i}\hbar}{2}\frac{\sqrt{-g}}{2}\int \mathrm{d}z\,\mathrm{d}y\ \frac{\delta^2 T_{\mu\nu}}{\delta\langle\phi(z)\rangle\delta\langle\phi(y)\rangle}\mathcal{G}_{\mathrm{m}}(z,y).
$$

A similar expression is obtained for the graviton loop if one replaces $T_{\mu\nu}$ by the energy–momentum tensor $t_{\mu\nu}$ for weak gravitational waves. Expanding the expression for the matter energy–momentum tensor in powers of $\phi-\langle\phi\rangle$ and taking its expectation value with respect to the matter state, one recognizes that the matter terms in (2.118) are equivalent to this expectation value. The same holds for the graviton state. In summary, in the case $J=0$ one gets from (2.118) the following semiclassical Einstein equations:

$$
R_{\mu\nu}-\frac{1}{2}g_{\mu\nu}R = 8\pi G\left(\langle T_{\mu\nu}\rangle+\langle t_{\mu\nu}\rangle\right), \qquad (2.121)
$$

where the left-hand side is evaluated for the mean metric $\langle g_{\mu\nu}\rangle$. This equation demonstrates again the fact that at one-loop order the gravitons appear at the same level as the matter fields; they both contribute to the right-hand side of the semiclassical Einstein field equations. Of course, if there is a 'macroscopic contribution' of the matter field, the graviton effect can be neglected and one arrives at (1.37). Equation (2.121) is the analogue of the Ehrenfest equations in quantum mechanics.

The expectation value on the right-hand side of (2.121) corresponds to the in–out expectation value (2.103), in contrast to (1.37). It can, however, be related to the ordinary expectation value; cf. Barvinsky and Nesterov (2001) and the references therein. There is another difference relative to (1.37): whereas there the metric is taken as classical from the outset, here the metric arises as the mean value of an underlying metric operator. Thus there is here no obvious discrepancy with the experiment by Page and Geilker (1981) discussed in Section 1.2.

We mention that the semiclassical Einstein equations are used as a starting point for the formalism of 'stochastic gravity', in which an approach to quantum gravity is attempted in the spirit of open-systems quantum theory (Hu and Verdaguer 2004).

If graviton effects can be neglected, the action $W_\phi[\bar g]$ defined by

$$
\mathrm{e}^{\mathrm{i}W_\phi[\bar g]} = \int \mathcal{D}\phi\ \mathrm{e}^{\mathrm{i}S_{\mathrm{m}}[\bar g,\phi]}, \qquad (2.122)
$$

where the mean field has been taken identical to the background field $\bar g$, is already the whole effective action.

For an evaluation of the effective action up to the one-loop order, one has to calculate an operator of the form $\ln \det D$; cf. (2.115). This can be done efficiently by using the 'Schwinger–DeWitt technique', which admits a covariant regularization (Birrell and Davies 1982; Fulling 1989; Barvinsky 1990; DeWitt 2003). This allows a local expansion of Green functions in powers of dimensional background quantities. Technically, a proper-time representation of the Green functions is used. This leads to a formulation of the effective action in terms of 'DeWitt coefficients' $a_0(x), a_1(x), a_2(x), \ldots,$ which are local scalars that are constructed from curvature invariants. (From $a_3(x)$ on, these coefficients are finite after the coincidence limit $y \to x$ is taken in the original version $a_3(x, y)$.) The effective action is thereby expanded in powers of the inverse curvature scale of the background. This only works in massive theories, although the divergent part of the action can also be used in the massless case. The Schwinger–DeWitt technique can thus be employed to compute UV divergences for massless theories too. The divergences arising can be absorbed into the gravitational constant, the cosmological constant, and c_1 and c_2. In this way, one can calculate the renormalized expectation value of the energy–momentum tensor, $\langle T_{\mu\nu} \rangle_{\rm ren}$, for the right-hand side of (2.121). The various methods of calculation are described in the above references. One then ends up at the one-loop level with the following renormalized form of the semiclassical Einstein equations:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} + c_1 H^{(1)}_{\mu\nu} + c_2 H^{(2)}_{\mu\nu} = 8\pi G \left(\langle T_{\mu\nu} \rangle_{\rm ren} + \langle t_{\mu\nu} \rangle_{\rm ren} \right), \tag{2.123}$$

where the four parameters G, Λ, c_1, and c_2 have to be determined experimentally (the latter two have already appeared in (2.97) above). The last two terms on the left-hand side are given as follows:

$$\begin{aligned} H^{(1)}_{\mu\nu} &= \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{\mu\nu}} \int \mathrm{d}^4 x \ \sqrt{-g} R^2 \\ &= 2R_{;\mu\nu} - 2g_{\mu\nu} \Box R - \frac{1}{2} g_{\mu\nu} R^2 + 2RR_{\mu\nu}, \end{aligned}$$

and

$$\begin{aligned} H^{(2)}_{\mu\nu} &= \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{\mu\nu}} \int \mathrm{d}^4 x \ \sqrt{-g} R_{\alpha\beta} R^{\alpha\beta} \\ &= 2R_{\mu;\nu\alpha}^{\ \ \ \ \ \alpha} - \Box R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \Box R + 2R_\mu^{\ \alpha} R_{\alpha\nu} - \frac{1}{2} g_{\mu\nu} R^{\alpha\beta} R_{\alpha\beta}. \end{aligned}$$

Usually, one cannot evaluate expressions such as (2.122) exactly. An example where this can be done is the case of a massless scalar field in two dimensions (see e.g. Vilkovisky 1984). The result is, reinserting $\hbar$,

$$W_\phi[\bar{g}] = -\frac{\hbar}{96\pi} \int \mathrm{d}^2 x \ \sqrt{-\bar{g}} \ {}^{(2)}\bar{R} \frac{1}{\Box} {}^{(2)}\bar{R}, \tag{2.124}$$

where ${}^{(2)}\bar{R}$ denotes the two-dimensional Ricci scalar. Equation (2.124) arises directly from the so-called 'Weyl anomaly' (also called the 'trace anomaly' or 'conformal anomaly'), that is, the breakdown of conformal invariance on quantization. Classically, this invariance leads to the vanishing of the energy–momentum tensor (cf. Section 3.2), while its breakdown in the quantum theory leads to a non-vanishing value. This anomaly plays a central role in string theory (Chapter 9).[29] It is given by

$$\langle T_\mu{}^\mu(x)\rangle_{\rm ren} = -\frac{a_1(x)\hbar}{4\pi} = -\frac{{}^{(2)}\bar{R}\hbar}{24\pi}. \tag{2.125}$$

In four space–time dimensions, the anomaly is proportional to the DeWitt coefficient $a_2(x)$ instead of $a_1(x)$.[30] In the two-dimensional model of (2.124), the flux of Hawking radiation (Section 7.1) is directly proportional to the anomaly.

We have seen that quantum GR is perturbatively non-renormalizable. In spite of this, we have argued that genuine predictions can be obtained through the method of effective action. Still, one can speculate that the occurrence of divergences could be cured automatically by going to a *non-perturbative* framework. One could imagine a theory that is perturbatively non-renormalizable, but that can be consistently defined at the non-perturbative level. One example is the Gross–Neveu model in three space–time dimensions (de Calan 1995).[31] This model describes a quantum field theory of N massless fermions coupled to each other by a four-fermion interaction term.

Another example is the non-linear σ model as given by an N-component field ϕ_a satisfying $\sum_a \phi_a^2 = 1$; see, for example, Hamber (2009), Sections 3.2 and 3.3. This model is non-renormalizable for space–time dimensions $n > 2$, but exhibits a non-trivial UV fixed point at some coupling $g_{\rm c}$, where it experiences a phase transition. Information about the behaviour in the vicinity of the non-trivial fixed point is obtained by an expansion in $n-2$ and use of renormalization-group techniques; compare the next subsection, on asymptotic safety. The non-linear σ model has even been tested in a space-shuttle experiment with superfluid helium, in which a specific-heat exponent was measured. This measured exponent was found to be in excellent agreement with theoretical calculations based on this model; see Hamber (2009) for a discussion.

We want to conclude this section by reviewing a simple example from Arnowitt *et al.* (1962) that shows how non-perturbative gravitational effects could in principle have this desirable consequence. The example is the self-energy of a thin charged shell. Assume that the shell has a 'bare' mass m_0, a charge Q, and a radius ϵ. At the Newtonian level (plus energy–mass equivalence from special relativity), the energy of the shell is given by

$$m(\epsilon) = m_0 + \frac{Q^2}{2\epsilon}, \tag{2.126}$$

which diverges in the limit $\epsilon \to 0$ of a point charge. The inclusion of the gravitational self-energy leads to

$$m(\epsilon) = m_0 + \frac{Q^2}{2\epsilon} - \frac{Gm_0^2}{2\epsilon}, \tag{2.127}$$

[29] The prefactor then reads $\hbar c/96\pi$, where c is the central charge; cf. (3.56).

[30] The anomaly occurs only for an even number n of space–time dimensions and is then proportional to $a_{n/2}$.

[31] However, no complete treatment is available in the published literature. I thank E. Seiler for discussions on this point.

which also diverges (unless one fine-tunes the charge unnaturally). Implementing, however, heuristically the (strong) equivalence principle by noting that gravity itself contributes to gravitational energy, one must replace the term m_0^2 in the above expression by $m^2(\epsilon)$:

$$m(\epsilon) = m_0 + \frac{Q^2}{2\epsilon} - \frac{Gm^2(\epsilon)}{2\epsilon}. \tag{2.128}$$

As $\epsilon \to 0$ this now has a finite limit,

$$m(\epsilon) \xrightarrow{\epsilon \to 0} \frac{|Q|}{\sqrt{G}}. \tag{2.129}$$

Since G appears in the denominator, this is a genuine non-perturbative result which cannot be found from any perturbation theory around $G = 0$.[32] The same result can be obtained within GR (Heusler *et al.* 1990). Such simple examples give rise to the hope that a consistent non-perturbative theory of quantum gravity will automatically prevent the occurrence of divergences; see, for example, also the models by DeWitt (1964) and Padmanabhan (1985). One of the main motivations of string theory (Chapter 9) is the avoidance of divergences in the first place. Genuine non-perturbative approaches are also provided by the canonical methods described in Chapters 3–6, where established quantization rules are applied to GR. Before we enter the discussion of these approaches, we shall address further non-perturbative approaches to quantum general relativity and shall then end with a brief review of quantum supergravity.

2.2.5 Asymptotic safety

As we have discussed in Sections 2.2.2 and 2.2.3, quantum gravity is perturbatively non-renormalizable. We have also seen that genuine effects can nevertheless be calculated at sufficiently low energies if one employs the concept of effective field theories. But what happens at high energies? There are two possibilities. One is that the quantization of GR is impossible and that one has to embark on a more general framework encompassing all interactions. This is the idea of string theory, to be discussed in Chapter 9. The other possibility is that quantum GR is non-perturbatively renormalizable. This could be established in various ways. One is by direct non-perturbative quantization of the Einstein–Hilbert action; this is attempted in the canonical approaches to be discussed in Chapters 4–6 and in the path-integral approach via Regge calculus or dynamical triangulation to be discussed in Section 2.2.6. Another way employs the notion of asymptotic safety, and this is the subject of the present subsection. Strictly speaking, it is not a direct quantization of GR, because more general actions than the Einstein–Hilbert action are mostly used. It is, however, close in spirit to a direct quantization, especially the effective-action approach discussed above.

The notion of asymptotic safety was introduced by Weinberg (1977, 1979) and is connected with the fact that coupling parameters in quantum field theory are energy-dependent due to renormalization.[33] A central ingredient is thus the use of

[32]Incidentally, $m = |Q|/\sqrt{G}$ is the mass–charge relation of an extremal Reissner–Nordström black hole; see Chapter 7. It is also the relation (1.10) between the Planck mass and the Planck charge.

[33]'... asymptotic safety can provide a rationale for picking physically acceptable quantum field theories, which may either *explain* renormalizability, or else *replace* it.' (Weinberg 1977, p. 34)

renormalization-group equations. A theory is said to be *asymptotically safe* if all essential coupling parameters g_i (these are the ones that are invariant under field redefinitions) approach, for energies $k \to \infty$, a fixed point where at least one of them does not vanish. The 'asymptotically' thus refers to the limit of high energies, and the 'safe' refers to the absence of singularities (i.e. the absence of Landau poles) in the coupling parameters. (In statistical physics, this corresponds to the occurrence of a phase transition.) We assume that the g_i are made dimensionless, that is,

$$g_i(k) = k^{-d_i}\bar{g}_i(k),$$

where $\bar{g}_i(k)$ are the original coupling parameters with mass dimensions d_i. A simple example of asymptotic safety is given by the Gross–Neveu model mentioned above (Braun *et al.* 2011).

In this approach, a central notion is played by the 'theory space'. This formal space is defined by the set of all action functionals that depend on a given set of fields and contain all terms that are consistent with a certain symmetry requirement; it is here where one makes close contact with the effective-theory idea. In the gravity context the symmetry is, of course, diffeomorphism invariance. One thus considers all actions $S[g_{\mu\nu}, \ldots]$ with this invariance, that is, actions containing terms $R/16\pi G$, $c_1 R^2$, $c_2 R_{\mu\nu}R^{\mu\nu}$, and so on. There are thus infinitely many coupling parameters $\bar{g}_i(k)$ given by $G, \Lambda, c_1, c_2, \ldots$ (all k-dependent, that is, 'running') and their dimensionless versions $g_i(k)$. They obey the *Callan–Symanzik* or generalized *Gell-Mann–Low* renormalization group equation,

$$k\partial_k g_i = \beta_i(g_1, g_2, \ldots) \; ; \tag{2.130}$$

cf. Weinberg (1996). A specific theory in theory space is then distinguished by a trajectory in the space of coupling parameters, where the trajectory is a solution of (2.130) with a particular initial condition. The surface formed by trajectories $g_i(k)$ that are attracted as $k \to \infty$ by a non-trivial fixed point g_{*i} is called the ultraviolet critical surface S_{UV}. (A non-trivial fixed point obeys $\beta_i(g_{*j}) = 0$ for all i for at least one $g_{*j} \neq 0$.) The dimensionality Δ_{UV} of S_{UV} thus gives the number of attractive directions and is equal to the number of free parameters of the theory. One expects that $\Delta_{\mathrm{UV}} < \infty$ in an asymptotically safe theory (Weinberg 1979), which means that all but a finite number of the g_i are fixed. The finitely many g_i that remain have to be fixed by experiment or astronomical observation. If Δ_{UV} were infinite-dimensional, infinitely many parameters would remain undetermined and one would be faced with the same situation as in a non-renormalizable theory—the loss of predictability. The optimal case would be $\Delta_{\mathrm{UV}} = 1$, because then only one free parameter would remain.

The dimension Δ_{UV} can be determined by expanding β_i around the fixed point,

$$\begin{aligned}\beta_i(g_j(k)) &= \underbrace{\beta_i(g_{*j})}_{=0} + \sum_j \frac{\partial \beta_i}{\partial g_j}(g_{*j})(g_j(k) - g_{*j}) + \ldots \\ &=: \sum_j B_{ij}(g_j(k) - g_{*j}) + \ldots .\end{aligned}$$

The general solution of (2.130) then reads

$$g_i(k) = g_{*i} + \sum_J C_J V_i^J \left(\frac{k_0}{k}\right)^{\theta_J}, \tag{2.131}$$

where the C_J are constants of integration, k_0 is a fixed reference scale, and

$$\sum_j B_{ij} V_j^J = -\theta_J V_i^J.$$

It is obvious that $\mathrm{Re}(\theta_J)$ must be positive in order for $g_i(k)$ to approach g_{*i} as $k \to \infty$. Thus, Δ_{UV} is equal to the number of eigenvalues of B_{ij} that obey this condition. The eigenvectors corresponding to these eigenvalues span the tangent space to the UV-critical surface at the non-trivial fixed point. The θ_J are called 'critical exponents'; they are invariant under reparametrizations (cf. p. 808 in Weinberg (1979)).

Here we investigate asymptotic safety in the framework of the effective average action Γ_k; see Berges *et al.* (2002) for a review. This notion can be traced back to the ideas of Kenneth Wilson on the coarse-grained version of the free-energy functional. It was first applied to gravity by Reuter (1998); see, for example, Percacci (2009) for a review. The action Γ_k is obtained by integrating out, in the full path integral, energies greater than k. It thus describes the physics at the scale k but with the effects of energies higher than k taken into account effectively. In the limit $k \to \infty$ one obtains the classical action from Γ_k, while in the opposite limit $k \to 0$ one obtains the conventional effective action Γ discussed in Section 2.2.4. The effective average action obeys (2.130) but with a complicated expression on the right-hand side (containing an integration over all momenta), see Wetterich (1993). This equation can thus only be handled if some truncation is employed, that is, if the flow of the renormalization group equation is projected onto a finite-dimensional subspace of theory space. The big question that arises is, of course, whether such a truncation will survive in the full theory.

The simplest truncation is the 'Einstein–Hilbert truncation': the theory space is fully described by the (Euclidean version of)[34] the Einstein–Hilbert action. In d dimensions it reads

$$\Gamma_k^{\mathrm{EH}}[g_{\mu\nu}] = \frac{1}{16\pi G_k} \int \mathrm{d}^d x\, \sqrt{g}(-R + 2\bar{\lambda}_k). \tag{2.132}$$

There are thus only two running (k-dependent) parameters: the gravitational constant G_k and the cosmological constant $\bar{\lambda}_k$. Their dimensionless versions are given by

$$\lambda_k = k^{-2}\bar{\lambda}_k\ , \quad g_k = k^{d-2} G_k,$$

respectively. Their behaviour with respect to k is explicitly obtained after insertion of (2.132) into (2.130). It follows that there exist both a trivial fixed point,[35] at

[34]The transition to the Lorentz signature is discussed in Manrique *et al.* (2011*a*); it was found there that the UV fixed points resemble the fixed points that are obtained for the Euclidean signature.

[35]A trivial fixed point is also called a 'Gaussian fixed point'; it corresponds to a free theory, and perturbation theory is only possible around a Gaussian fixed point. Asymptotic safety at a trivial fixed point corresponds to asymptotic freedom. Asymptotic safety at a non-trivial fixed point is a genuinely non-perturbative statement.

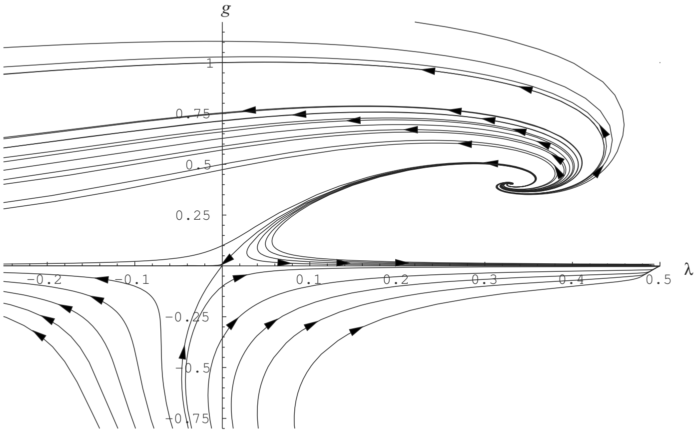

Fig. 2.4 Flow of the dimensionless gravitational constant g_k and cosmological constant λ_k. There is an infrared Gaussian fixed point at the origin and an ultraviolet non-Gaussian fixed point in the first quadrant. Reprinted with kind permission from Reuter and Saueressig (2002*a*). © 2002 by the American Physical Society.

$\lambda_* = 0 = g_*$, and a non-trivial one, at $\lambda_* \neq 0$, $g_* \neq 0$; see Reuter (1998) and Lauscher and Reuter (2001). This is a first indication that four-dimensional quantum gravity may be asymptotically safe. Therefore, as $k \to \infty$,

$$G_k \approx \frac{g_*}{k^{d-2}} \;, \quad \bar{\lambda}_k \approx \lambda_* k^2. \tag{2.133}$$

The gravitational constant thus vanishes (in $d > 2$ dimensions) as $k \to \infty$, which would correspond to a coupling that is asymptotically free (such as the gauge coupling in QCD). Figure 2.4 shows the flow of the dimensionless gravitational and coupling constants found from the Einstein–Hilbert truncation (Reuter and Saueressig 2002*a*).

In order to test the reliability of the Einstein–Hilbert truncation, one can investigate various generalizations of the truncation ansatz. For example, one can consider instead of (2.132) an improved version of the form[36]

$$\Gamma_k^{(1)} = \Gamma_k^{\mathrm{EH}} + \int \mathrm{d}^d x \, \sqrt{g} c_{1k} R^2(g) \;; \tag{2.134}$$

cf. Lauscher and Reuter (2002). A class of non-local truncations was considered in Reuter and Saueressig (2002*b*). Although there exists no general proof, there is evidence that the non-trivial fixed point survives within the class of truncations considered; cf. Benedetti *et al.* (2010) and the references therein. This suggests that asymptotic safety may indeed hold for the full theory. Background independence is implemented in the asymptotic-safety approach in a manner similar to DeWitt's background-field method discussed in Section 2.2.3 (Manrique *et al.* (2011*b*).

[36]This should not be confused with the perturbative expansion in Section 2.2.3, where such terms also appear.

Even at this level, a number of intriguing results can be obtained. For the classical dimensionality $d = 4$, one finds from the asymptotic form of the graviton propagator that in the vicinity of the non-trivial fixed point, corresponding to scales $l \approx l_{\rm P}$, space–time appears to be effectively two-dimensional (Lauscher and Reuter 2001). For large scales, $l \gg l_{\rm P}$, one gets four dimensions, as expected. This behaviour is also found in the dynamical-triangulation approach discussed in the next subsection.

If quantum GR is asymptotically safe in the sense discussed here, a number of interesting cosmological conclusions can be drawn (Reuter and Weyer 2004). The gravitational 'constant' may actually grow with increasing distance. This effect could be observable on the scale of galaxies and clusters of galaxies and could mimic the existence of 'dark matter'. Moreover, there is strong evidence that a small positive cosmological constant is found as a strong infrared quantum effect. This could give an explanation of 'dark energy'. It is thus imaginable that two of the most fundamental puzzles in astrophysics may be solved by a macroscopic quantum effect of gravity. Other potential cosmological applications include the emergence of an inflationary phase in the early Universe without an inflaton field, and the generation of cosmic entropy (Bonanno and Reuter 2011).

If gravity is asymptotically safe, it is imaginable that the Standard Model of electroweak interactions *plus* gravity is valid up to arbitrarily high energies; that is, it could be a candidate for an exact theory and not just an effective theory (Shaposhnikov and Wetterich 2010). This could be possible if gravity were coupled non-minimally to the Higgs field, a possibility that could also establish the Higgs particle as the inflaton needed for a successful inflationary scenario of the early universe (see e.g. Barvinsky *et al.* 2009). On the basis of this assumption, predictions of the Higgs mass can be made (Shaposhnikov and Wetterich 2010).

2.2.6 Regge calculus and dynamical triangulation

Quantum-gravitational path integrals in four dimensions cannot be evaluated analytically without making approximations (such as saddle-point approximations). It is thus understandable that attempts are being undertaken to make them amenable through numerical methods. As has already been mentioned, in the so-called quantum *Regge calculus* one considers the Euclidean path integral and decomposes a four-dimensional configuration into a set of simplices; see Regge and Williams (2000) and Hamber (2009) for reviews. The latter reference also discusses at great length the connection of the lattice formulation to the $2 + \epsilon$-expansion, which is an expansion around space–time dimension two, where the gravitational constant is dimensionless (and where pure gravity is just a topological theory).

In Regge calculus, the edge lengths are treated as dynamical entities. An important feature of the calculation is the implementation of the triangle inequality for these lengths—the formalism must implement the fact that the length of one side of a triangle is smaller than the sum of the other two. The need for this hinders an evaluation of the path integral in the Regge framework by means other than numerically.

An alternative method, therefore, is to keep the edge lengths fixed and to perform the sum in the path integral over all possible manifold-gluings of simplices, reducing the evaluation to a combinatorial problem. This method is called *dynamical trian-*

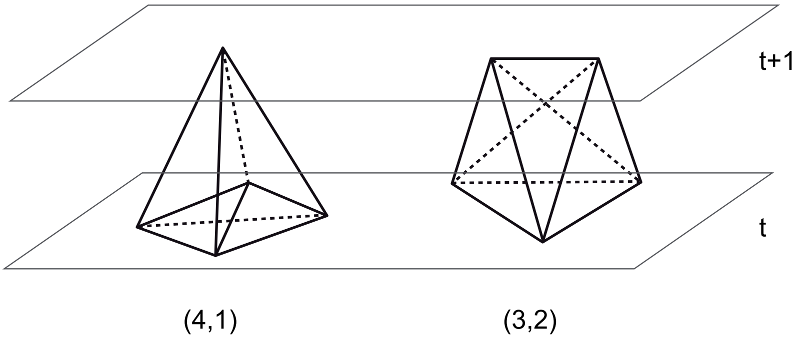

Fig. 2.5 The two fundamental simplices used as building blocks in four-dimensional dynamical triangulation. The surfaces t = constant and $t+1$ = constant represent two adjacent three-dimensional spaces. The notation (a, b) means a vertices at time t and b vertices at time $t+1$. Reprinted with kind permission from Ambjørn *et al.* (2005). © 2005 by the American Physical Society.

gulation; see Ambjørn *et al.* (2010) for a detailed review. If one again considers the Euclidean path integral, one encounters problems. First, there is the conformal-factor problem (Section 2.2.1). Second, the sum over configurations does not generate a four-dimensional geometry in the macroscopic limit; there is either a 'polymerization' (occurrence of an effective dimension around two) or the generation of geometries with a very large dimension at large scales. For this reason, Ambjørn and Loll (1998) have introduced a Lorentzian version of dynamical triangulation. This has the advantage that the causal (lightcone) structure of the space–time configurations is directly implemented in the path integral. The branch points of the Euclidean approach are avoided and there is no change of spatial topology. This gives rise to differences from the Euclidean approach, notably the occurrence of an effective four-dimensional geometry at large scales.

In the following, we shall give a brief introduction to Lorentzian dynamical triangulation. The building blocks are the two four-dimensional simplices shown in Fig. 2.5 (plus their time-reversed versions). They are completely described by their fixed squared edge lengths $\{l_i^2\}$, where l_i^2 is equal to a^2 (space-like case) or $-\alpha a^2$, where $\alpha > 0$ is a free parameter (time-like case). Here, a is a regularization parameter with the dimension of a length, which should go to zero in the continuum limit. Because of their Lorentzian nature, space-like and time-like edge lengths will in general not be equal. The space of all gluings of such simplices will then be identified with the space of all geometries.

Curvature is implemented in this approach as a sum over the Gaussian curvatures of all two-dimensional submanifolds, where the deficit angle at a vertex is a measure of the Gaussian curvature there.

The path integral (2.72) is then defined by the following identification:

$$Z = \int \mathcal{D}g_{\mu\nu}(x)\ \mathrm{e}^{\mathrm{i}S[g_{\mu\nu}(x)]} \longrightarrow \sum_{T\in\mathcal{T}} \frac{1}{C_T} \mathrm{e}^{\mathrm{i}S_{\mathrm{discrete}}(T)}, \tag{2.135}$$

where T denotes inequivalent gluings from the class $\mathcal{T}$ of triangulations, and C_T is a symmetry factor that is almost always equal to one for large triangulations. For technical reasons, one has still to perform a Wick rotation in (2.135), $t \to -\mathrm{i}\tau$ (mapping time-like links into space-like links). After this Wick rotation into the Euclidean regime, the action S_{discrete} (also called the 'Regge action') is given by the expression

$$S_{\text{discrete}}(T) = \kappa_{d-2} N_{d-2} - \kappa_d N_d, \tag{2.136}$$

where N_d is the number of d-dimensional simplices (which is proportional to the total volume), and the κ_i are functions of the bare gravitational and cosmological constants. Inserting the expression (2.136) into (2.135), one recognizes that there are two separate sums, one over d-dimensional simplices (coming from the volume term, that is, the second term in (2.136)) and one over $d-2$-dimensional simplices (coming from simplices of co-dimension two because the first term in (2.136) is the curvature term, and curvature is calculated from the Gaussian curvatures of two-dimensional submanifolds).

At the end of the calculation, one would like to perform an inverse Wick rotation back to the Lorentzian regime. This is explicitly possible only in $d = 2$. In higher dimensions, one typically considers observables that are invariant under the Wick rotation and thus have the same value in the Euclidean and the Lorentzian domain. It should be emphasized that the formalism is Lorentzian in the first place and that the Wick rotation to the Euclidean regime (and back) is only applied at an intermediate stage for technical reasons.

For the continuum limit, one takes $N_d \to \infty$ and $a \to 0$ while holding some appropriate quantity fixed (such as the average total volume $\propto N_d a^d$ of all simplices together). In the course of this process, one obtains a positive cosmological constant. While this is in agreement with current observations, no numerical value can be predicted for it.

Whereas the continuum limit cannot be obtained analytically in the four-dimensional case, this is possible fully for $d = 2$ and partially for $d = 3$. Let us consider the two-dimensional case, which is illustrative and completely soluble. Only one type of building block exists here: an equilateral Minkowskian triangle with squared lengths $-a^2$ and a^2, respectively. Periodic boundary conditions are chosen for space (which is thus topologically a circle, S^1). Since the curvature term in $S_{\text{discrete}}(T)$ is a topological invariant, it drops out of the sum in (2.135) (the topology is fixed). Thus, only the second term in (2.136), corresponding to the cosmological-constant term, is present. The sum converges if the bare cosmological constant is greater than some critical (positive) value. For the path integral (2.135), one obtains the expression

$$Z = \exp\left(-\coth(\mathrm{i}\sqrt{\Lambda}t)\sqrt{\Lambda}(l_{\text{in}} + l_{\text{out}})\right) \frac{\sqrt{\Lambda l_{\text{in}} l_{\text{out}}}}{\sinh(\mathrm{i}\sqrt{\Lambda}t)} I_1\left(\frac{2\sqrt{\Lambda l_{\text{in}} l_{\text{out}}}}{\sinh(\mathrm{i}\sqrt{\Lambda}t)}\right), \tag{2.137}$$

where I_1 denotes a Bessel function, and l_{in} and l_{out} denote the sizes of the initial and final one-geometries, respectively (the circumferences of the S^1); t is the proper-time distance between the initial and the final one-geometry, and Λ is the renormalized cosmological constant. The cosmological constant is the only dimensionful parameter here.[37] Since pure GR in two dimensions is trivial (the Einstein tensor vanishes identically in this case), the result obtained here corresponds to a pure quantum theory without a classical limit. As mentioned above, the topology is fixed here. A non-perturbative implementation of a sum over topologies in two dimensions is discussed in Ambjørn *et al.* (2010). An analytic result such as (2.137) is impossible to obtain in the quantum Regge calculus.

The result (2.137) still contains the proper-time distance t. Strictly speaking, one has to perform in addition a sum over all t in order to obtain the full quantum-gravitational path integral; see Kiefer (1991) and Section 5.3.4.

For the two-dimensional theory, the quantum effective Hamiltonian has been calculated. It is self-adjoint and possesses a discrete spectrum given by

$$E_n = 2(n+1)\sqrt{\Lambda}\ , \quad n \in \mathbb{N}. \tag{2.138}$$

A useful notion in all dimensions is the effective Hausdorff dimension, $d_{\rm H}$. It is defined by

$$\langle V(R)\rangle \propto R^{d_{\rm H}},$$

where $V(R)$ denotes the volume of a geodesic ball of radius R as a function of the distance R. We emphasize that $d_{\rm H}$ is a dynamical quantity and need not coincide with the dimension d of the building blocks. Interestingly, for $d = 2$ one obtains in fact $d_{\rm H} = 2$ from the Lorentzian path integral, whereas the Euclidean approach yields $d_{\rm H} = 4$. What about the coupling to matter? For $d = 2$, various copies of Ising models have been included. It turns out that the critical matter exponents are, to within the experimental measurement accuracy, approximately equal to the Onsager exponents in the standard Ising case; that is, the behaviour is as if the spins lived on a static flat lattice (for the Euclidean case, this does not hold).

In the three-dimensional case, the evaluation of the sum (2.135) is still possible, but one now has to employ a Monte Carlo simulation (with the volume of space–time approximately kept fixed). Of course, Monte Carlo methods have also to be used for the most interesting case $d = 4$. Although the continuum limit has not been obtained analytically, there exists strong numerical evidence that it exists, and a number of interesting results are available (Ambjørn *et al.* 2005). Depending on the bare gravitational constant and an asymmetry parameter Δ (measuring the difference in the sizes of the space-like and time-like edge lengths), one finds three phases in $d = 4$. In one of them (for sufficiently large coupling and Δ) there is strong evidence that the observed four-dimensionality of macroscopic space–times emerges. This is different from the result of the corresponding Euclidean approach. Various notions of dimensionality have been employed in discussions of this topic in the literature: the Hausdorff dimension and other notions (scaling and spectral dimensions). Of particular interest is the *spectral dimension* $d_{\rm s}$, which is defined by investigating diffusion processes on manifolds. In particular, one has for the average return probability in such a process the relation

$$P(\mathbf{x},\mathbf{x};s) \sim (4\pi s)^{-d_{\rm s}/2},$$

[37]In fact, one has the combination $\Lambda/G\hbar$ instead of Λ, but $G\hbar$ is dimensionless in $d = 2$ and has therefore been set equal to a constant.

where s is the diffusion time. If a diffusion process exists on the manifold, one can then use this relation to *define* the spectral dimension.

On small scales, space–time seems to assume only two dimensions—a highly non-classical feature. This result is surprisingly similar to the result obtained from the renormalization-group flow discussed in the last subsection; cf. Carlip (2010). An effective dimension of two is also found from the area spectrum in loop quantum gravity, see Chapter 6. In looking for a continuum theory that could fully reproduce the behaviour of dynamical triangulations, Sotiriou *et al.* (2011) found evidence that this continuum theory could be related to the model of Hořava–Lifshits gravity mentioned in Section 2.2.2. The evidence arises by comparing the flow of the spectral dimension in Hořava–Lifshits gravity with that obtained from dynamical triangulation. There could thus be a fundamental link between various apparently different approaches to quantum gravity.

Another interesting result concerns the form of the 'minisuperspace effective action'.[38] This has the same form as the action found from a direct minisuperspace approximation, but is now derived from a path integral that takes into account 'all' configurations, not only those respecting the symmetries. This puts quantum cosmology (Chapter 8) on a firmer footing.

Besides Regge calculus and dynamical triangulation, there are alternative approaches. One example is lattice spinor gravity, in which Grassmann variables are used as the basic degrees of freedom (Diakonov 2011, Wetterich 2011). The metric and the tetrad then arise as an expectation value of a suitable collective field. Loosely speaking, one could talk here about 'composite gravitons'.

Regarding this and the last subsection, one can say that non-perturbative aspects of covariant quantum general relativity can be discussed on a rigorous footing, with intriguing results, in spite of the formal non-renormalizability of the perturbation theory.

2.3 Quantum supergravity

2.3.1 General introduction

Supergravity (SUGRA) is a supersymmetric theory of gravity encompassing GR. Supersymmetry (SUSY) is a symmetry which mediates between bosons and fermions. It exhibits interesting features; for example, the running coupling constants in the Standard Model of particle physics can meet at an energy of around 10^{16} GeV if SUSY is added. SUGRA is a theory in its own right; see, for example, van Nieuwenhuizen (1981) for a review. Our main question of concern is whether the perturbative UV behaviour of quantum gravity discussed in the last section can be improved by going over to SUGRA. Before we address this question in the next subsection, some introductory remarks are in order.

SUSY arose from the question of whether the Poincaré group (and therefore space–time symmetries) can be unified with an internal (compact) group such as SO(3). A no-go theorem states that in a relativistic quantum field theory, given 'natural' assumptions of locality, causality, positive energy, and a finite number of elementary

[38] Minisuperspace is the truncation of the geometries to highly symmetric spaces; cf. Chapter 8.

particles, such an invariance group can only be the direct product of the Poincaré group with a compact group, preventing a real unification. There is, however, a loophole. A true unification is possible if *anti*commutators are used instead of commutators in the formulation of a symmetry, leading to a 'graded Lie algebra'.[39] It was shown by Haag *et al.* (1975) that, with the above assumptions of locality etc., the algebraic structure is essentially unique.

The SUSY algebra is given by the anticommutator

$$[Q^i_\alpha, \bar{Q}^j_\beta]_+ = 2\delta^{ij}(\gamma^n)_{\alpha\beta}P_n \ , \ i,j = 1,\ldots,N, \tag{2.139}$$

where Q^i_α denotes the corresponding generators, also called spinorial charges; $\bar{Q}^i_\alpha = Q^i_\alpha\gamma^0$, with γ^0 being one of Dirac's gamma matrices; N is the number of SUSY generators; and all anticommutators among the Qs and the $\bar{Q}$s themselves vanish. There are also the commutators

$$[P_n, Q^i_\alpha] = 0, \ [P_m, P_n] = 0. \tag{2.140}$$

(P_n denotes the energy–momentum four-vector, which is the generator of space–time translations.) In addition, there are the remaining commutators of the Poincaré group, (2.34)–(2.36), as well as

$$[Q^i_\alpha, J_{mn}] = (\sigma_{mn})^\beta_\alpha Q^i_\beta, \tag{2.141}$$

with $\sigma_{mn} = \mathrm{i}[\gamma_m, \gamma_n]$; cf. also Section 1.1. More details can be found, for example, in Weinberg (2000). The SUSY algebra is compatible with relativistic quantum field theory; that is, one can write the spinorial charges as an integral over a conserved current,

$$Q^i_\alpha = \int \mathrm{d}^3x \ J^i_{0\alpha}(x) \ , \quad \frac{\partial J^i_{m\alpha}(x)}{\partial x^m} = 0. \tag{2.142}$$

Fermions and bosons are combined into 'supermultiplets' by irreducible representations of this algebra. There is a fermionic superpartner to each boson and vice versa. One would thus expect the partners to have the same mass. Since this is not observed in nature, SUSY must be broken. The presence of SUSY would guarantee that there are equal numbers of bosonic and fermionic degrees of freedom. For this reason, several divergences would cancel due to the presence of opposite signs (e.g. the 'vacuum energy'). This gave rise to the hope that SUSY might generally improve the UV behaviour of quantum field theories; see the next subsection.

Performing now an independent SUSY transformation at each space–time point, one arrives at a corresponding *gauge symmetry*. Because the anticommutator (2.139) of two SUSY generators closes on the space–time momentum, this means that space–time translations are performed independently at each space–time point—they are nothing but general coordinate transformations. The gauge theory therefore contains GR and is called SUGRA.[40] For each generator there is then a corresponding gauge field: P_n

[39] Anticommutators were, of course, used before the advent of SUSY in order to describe fermions, but not in the context of symmetries.

[40] In fact, the gauging of the Poincaré group leads to the Einstein–Cartan theory, which besides curvature also contains torsion.

corresponds to the vierbein field e_μ^n (see Section 1.1), J_{mn} to the 'spin connection' ω_μ^{mn}, and Q_α^i to the 'Rarita–Schwinger fields' $\psi_\mu^{\alpha,i}$. The latter are fields with spin 3/2 and describe the fermionic super-partners to the graviton—the *gravitinos*. They are a priori massless, but can acquire a mass by a Higgs mechanism. For $N = 1$ (simple SUGRA), one has a single gravitino, which sits together with the spin-2 graviton in one multiplet. The cases $N > 1$ are referred to as 'extended supergravities'. In the case $N = 2$, for example, the photon, the graviton, and two gravitinos together form one multiplet, yielding a 'unified' theory of gravity and electromagnetism. One demands that $0 \leq N \leq 8$ because otherwise there would be more than one graviton and also particles with spin higher than two (for which no satisfactory coupling exists).

For $N = 1$, the SUGRA action is the sum of the Einstein–Hilbert action and the Rarita–Schwinger action for the gravitino,

$$S = \frac{1}{16\pi G}\int \mathrm{d}^4x\ (\det e_\mu^n)R + \frac{1}{2}\int \mathrm{d}^4x\ \epsilon^{\mu\nu\rho\sigma}\bar{\psi}_\mu\gamma_5\gamma_\nu D_\rho\psi_\sigma \tag{2.143}$$

(recall that $\det e_\mu^n = \sqrt{-g}$, and $\gamma_5 = \mathrm{i}\gamma^0\gamma^1\gamma^2\gamma^3$). In (2.143), we have introduced the spinorial covariant derivative

$$D_\mu = \partial_\mu - \tfrac{1}{2}\,\omega_\mu^{nm}\sigma_{nm}$$

(see (1.22)), and set $\Lambda = 0$. The action (2.143) is invariant not only under general coordinate transformations and local Poincaré transformations, but also under local SUSY transformations, which for the vierbein and gravitino fields read

$$\begin{aligned} \delta e_\mu^m &= \tfrac{1}{2}\,\sqrt{8\pi G}\bar{\epsilon}^\alpha\gamma^m_{\alpha\beta}\psi_\mu^\beta, \\ \delta\psi_\mu^\alpha &= \frac{1}{\sqrt{8\pi G}}D_\mu\epsilon^\alpha, \end{aligned} \tag{2.144}$$

where ϵ^α is an anticommuting parameter function and $\bar{\epsilon}^\alpha$ its complex conjugate. Note that the factors $\sqrt{G}$ are needed for dimensional reasons.

The big open question is, of course, whether SUSY is realized in nature. Since SUSY plays a crucial role in the construction of superstring theory (Chapter 9), the answer to this question will also decide the fate of string theory. One of the biggest hopes connected with the Large Hadron Collider (LHC) in Geneva is the discovery of SUSY at the TeV scale, where it is supposed to appear if it plays a role in the unification of interactions. The first experimental results do not look optimistic (Strumia 2011): only about 0.7% of the parameter space of the constrained minimal supersymmetric Standard Model seems to survive. But the final decision has not yet been made.

2.3.2 Infinities in perturbative quantum supergravity

We have seen in Section 2.2.3 that perturbative quantum gravity exhibits the divergences that are expected from the power-counting arguments for non-renormalizability. Does the introduction of SUSY improve this situation? After all, fermions contribute with a relative minus sign compared to bosons in each closed Feynman loop, and so it is possible that they might cancel the unwanted infinities. As de Wit and Nicolai (1982) put it, 'The balanced decomposition in bosons and fermions has a softening effect on its quantum divergences, thus offering hopes for a consistent quantum theory of gravity . . . ' Due to its higher symmetry, divergences are expected in SUGRA from three loops on (if there is no cancellation). In fact, the theories for $N < 8$ exhibit the expected divergences. A special role is played by $N = 8$ SUGRA. As mentioned above, $N = 8$ is the maximal number of SUSY generators. The theory contains an irreducible multiplet that consists of massless states including the spin-2 graviton, eight spin-3/2 gravitinos, 28 spin-1 states, 56 spin-1/2 states, and 70 spin-0 states; see, for example, de Wit and Nicolai (1982) for details.

For a long time, it seemed that the enormous amount of complexity of calculating the divergences would forever forbid an explicit calculation. The situation has, however, changed. Bern *et al.* (2009) explicitly calculated the complete four-loop four-particle amplitude of $N = 8$ SUGRA. They found *ultraviolet finiteness* and also presented arguments that the perturbation theory is finite at five and six loops (finiteness at three loops had been found earlier by that group). This allows the speculation that the theory is finite at all orders. If this were true, $N = 8$ SUGRA would be a perturbatively consistent theory of quantum gravity. An explicit proof would, however, need some further insight into the structure of this theory; for example, the discovery of a hitherto unknown symmetry.

Bern *et al.* (2009) were able to present their results only by using new methods. These methods arose as the offspring of string calculations, because quantum-field-theoretic amplitudes can arise as special limits of string amplitudes for which explicit calculations are possible. The relation to quantum gravity was established by the observation that there is a correspondence between $N = 8$ SUGRA and $N = 4$ super-Yang–Mills theory, which is the maximally supersymmetric Yang–Mills theory and is known to be finite. Instead of the usual Feynman diagrams, the method employed leads to a new type of diagrams called 'Mondrian diagrams' because they resemble the work of the artist Piet Mondrian.

These investigations show that a finite perturbation theory of quantum gravity might be possible. The alternative is to construct either a full non-perturbative theory of the quantized gravitational field or a unified theory of all interactions beyond field theory. The following chapters are devoted to these directions.

Further reading: Barvinsky and Vilkovisky (1987), Hamber (2009), Woodard (2009).

3 Parametrized and relational systems

In this chapter, we shall consider some models that exhibit certain features of GR but which are much easier to discuss. In this sense, they constitute an important conceptual preparation for the canonical quantization of GR, which is the topic of the next few chapters. In addition, they are of interest in their own right.

The central aspect is *reparametrization invariance* and the ensuing existence of constraints; see, for example, Sundermeyer (1982) and Henneaux and Teitelboim (1992) for a general introduction to constrained systems. Kuchař (1973) gives a detailed discussion of reparametrization-invariant systems, which we shall partly follow in this chapter. Such invariance properties are often referred to as 'general covariance' of the system because they refer to an invariance with respect to a relabelling of the underlying space–time manifold. A more precise formulation has been suggested by Anderson (1967), proposing that an invariance group is a subgroup of the full covariance group which leaves the absolute, non-dynamical, elements of a theory invariant; cf. also Ehlers (1995) and Giulini (2007). Such an absolute element could, for example, be the conformal structure in the scalar theory of gravity mentioned at the end of Section 2.1.1. In GR, the full metric is dynamical and the invariance group coincides with the covariance group, the group of all diffeomorphisms. According to Anderson (1967), general covariance should be interpreted as the absence of absolute structure, also called 'background independence'.

Whereas dynamical elements are subject to quantization, absolute elements remain classical; see Section 1.3. Absolute elements can also appear in 'disguised form' if a theory has been made reparametrization invariant in an artificial way. This is the case in the models of the non-relativistic particle and the parametrized field theory to be discussed below, but not in GR or the other dynamical systems considered in this chapter.

3.1 Particle systems

3.1.1 Parametrized non-relativistic particle

Let us consider the action for a point particle in classical mechanics,

$$S[q(t)] = \int_{t_1}^{t_2} \mathrm{d}t\ L\left(q, \frac{\mathrm{d}q}{\mathrm{d}t}\right). \tag{3.1}$$

It is only for simplicity that a restriction to one particle is being made. The following discussion can be easily generalized to n particles. For simplicity, the Lagrangian in (3.1) has been chosen t-independent.

We introduce now a formal time parameter τ ('label time') and elevate t (Newton's 'absolute time') formally to the rank of a dynamical variable (cf. Dirac 1933). This is an example of an absolute structure in disguise, as mentioned above. We therefore write $q(\tau)$ and $t(\tau)$. Derivatives with respect to τ will be denoted by a dot, and restriction to $\dot{t} > 0$ is made. The action (3.1) can then be rewritten as

$$S[q(\tau), t(\tau)] = \int_{\tau_1}^{\tau_2} \mathrm{d}\tau \ \dot{t} L\left(q, \frac{\dot{q}}{\dot{t}}\right) =: \int_{\tau_1}^{\tau_2} \mathrm{d}\tau \ \tilde{L}(q, \dot{q}, \dot{t}). \tag{3.2}$$

The Lagrangian $\tilde{L}$ possesses the important property that it is *homogeneous* (of degree one) in the velocities; that is,

$$\tilde{L}(q, \lambda\dot{q}, \lambda\dot{t}) = \lambda\tilde{L}(q, \dot{q}, \dot{t}), \tag{3.3}$$

where $\lambda \neq 0$ can be an arbitrary function of τ. Homogeneous Lagrangians lead to actions that are invariant under time reparametrizations $\tau \to \tilde{\tau} = f(\tau)$ in the sense that they can be written as a $\tilde{\tau}$-integral over the same Lagrangian, depending now on $\mathrm{d}q/\mathrm{d}\tilde{\tau}$. Assuming $\dot{f} > 0$ gives

$$S = \int_{\tau_1}^{\tau_2} \mathrm{d}\tau \ L(q, \dot{q}) = \int_{\tilde{\tau}_1}^{\tilde{\tau}_2} \frac{\mathrm{d}\tilde{\tau}}{\dot{f}} \ L\left(q, \frac{\mathrm{d}q}{\mathrm{d}\tilde{\tau}}\dot{f}\right) = \int_{\tilde{\tau}_1}^{\tilde{\tau}_2} \mathrm{d}\tilde{\tau} \ L\left(q, \frac{\mathrm{d}q}{\mathrm{d}\tilde{\tau}}\right). \tag{3.4}$$

The canonical momentum for q is found from (3.2) to read

$$\tilde{p}_q = \frac{\partial\tilde{L}}{\partial\dot{q}} = \dot{t}\frac{\partial L}{\partial\left(\dot{q}/\dot{t}\right)}\frac{1}{\dot{t}} = p_q, \tag{3.5}$$

thus coinciding with the momentum corresponding to (3.1). But now there is also a momentum canonically conjugate to t,

$$\begin{aligned} p_t = \frac{\partial\tilde{L}}{\partial\dot{t}} &= L\left(q, \frac{\dot{q}}{\dot{t}}\right) + \dot{t}\frac{\partial L\left(q, \dot{q}/\dot{t}\right)}{\partial\dot{t}} \\ &= L\left(q, \frac{\mathrm{d}q}{\mathrm{d}t}\right) - \frac{\mathrm{d}q}{\mathrm{d}t}\frac{\partial L(q, \mathrm{d}q/\mathrm{d}t)}{\partial(\mathrm{d}q/\mathrm{d}t)} = -H. \end{aligned} \tag{3.6}$$

Therefore, t and $-H$ (the negative of the Hamiltonian corresponding to the original action (3.1)) are canonically conjugate pairs. The Hamiltonian belonging to $\tilde{L}$ is found as

$$\tilde{H} = \tilde{p}_q\dot{q} + p_t\dot{t} - \tilde{L} = \dot{t}(H + p_t). \tag{3.7}$$

But, because of (3.6), this is constrained to vanish. It is appropriate at this stage to introduce a new quantity called the 'super-Hamiltonian'. This is defined as

$$H_{\mathrm{S}} := H + p_t, \tag{3.8}$$

and one has the *constraint*

$$H_{\mathrm{S}} \approx 0. \tag{3.9}$$

The $\approx$ in this and further equation(s) means 'to vanish as a constraint' or 'weak equality' in the sense of Dirac; see Section 3.5. It defines a subspace in phase space and can be set to zero only *after* all Poisson brackets have been evaluated. One can now use instead of (3.1) the new action principle

$$S = \int_{\tau_1}^{\tau_2} \mathrm{d}\tau \ (p_q\dot{q} + p_t\dot{t} - NH_{\mathrm{S}}), \tag{3.10}$$

where all quantities (including N) have to be varied; N is a Lagrange multiplier, and variation with respect to it just yields the constraint (3.9). From Hamilton's equations, one has

$$\dot{t} = \frac{\partial(NH_{\mathrm{S}})}{\partial p_t} = N. \tag{3.11}$$

Therefore, N is called the *lapse function* because it gives the rate of change of Newton's time t with respect to the label time τ.

The existence of the constraint (3.9) is a consequence of the reparametrization invariance with respect to τ. To see this explicitly, it will be proven that having a Lagrangian homogeneous in the velocities is equivalent to the corresponding Hamiltonian being zero. Homogeneity has been shown above to be equivalent to reparametrization invariance.

Given a homogeneous Lagrangian, one finds for the canonical Hamiltonian

$$H_{\mathrm{c}} = \frac{\partial L}{\partial \dot{q}}\dot{q} - L = \lambda^{-1}\left(\frac{\partial L(q,\lambda\dot{q})}{\partial(\lambda\dot{q})}\lambda\dot{q} - L(q,\lambda\dot{q})\right) = \lambda^{-1}H_{\mathrm{c}}.$$

Since λ is arbitrary, H_{c} must vanish. On the other hand, if H_{c} vanishes, one gets the following, after substituting $\dot{q}$ by $\lambda\dot{q}$:

$$\frac{\partial L(q,\lambda\dot{q})}{\partial \dot{q}}\dot{q} = L(q,\lambda\dot{q}).$$

The left-hand side can be written as $\lambda p_q\dot{q} = \lambda L$, and one gets $\lambda L(q,\dot{q}) = L(q,\lambda\dot{q})$, that is, L is homogeneous.

One can have reparametrization invariance without a Hamiltonian constraint if the q and p do not transform as scalars under reparametrizations (Henneaux and Teitelboim 1992). This is not, however, a natural situation. In this case, the theorem just proven remains true, but the connection to reparametrization invariance is lost.

Although Newton's time has been mixed amongst the other dynamical variables, it can easily be recovered, for its momentum p_t enters *linearly* into (3.8) (it is assumed that H has the usual form $H = p_q^2/2m + V(q)$). Therefore, one can easily solve (3.9) to find $p_t = -H$, choose the label $\tau = t$ ('fixing the gauge'), and find from (3.10)

$$S = \int \mathrm{d}t \ \left(p_q\frac{\mathrm{d}q}{\mathrm{d}t} - H\right), \tag{3.12}$$

that is, just the standard action (the Hamiltonian form of (3.1)). This process is called *deparametrization*. We shall see in Chapter 4 that there is an analogue to (3.8) and (3.9) in GR, but that there, in contrast to here, all momenta occur quadratically. This leads to the interesting question whether a deparametrization for GR, that is, the identification of a distinguished time-like variable, is possible. It should be remarked that every system can be transformed artificially into 'generally covariant' form; cf. Kretschmann (1917). But this is possible only at the price of disguising absolute structures which then formally appear as dynamical variables, such as Newton's absolute time t. The general covariance of GR is natural in the sense that the metric is fully dynamical.

How does one quantize a system given by a constraint such as (3.9)? A successful, although heuristic, procedure is the proposal made by Dirac (1964). A classical constraint is implemented in the quantum theory as a restriction on physically allowed wave functions. Thus, (3.9) is translated into

$$\hat{H}_{\mathrm{S}}\psi = 0, \tag{3.13}$$

where $\hat{H}_{\mathrm{S}}$ denotes the super-Hamilton operator associated with the classical super-Hamiltonian H_{S}. In the position representation, the $\hat{q}$ are represented by multiplication by q and the momenta $\hat{p}$ are represented by derivatives $(\hbar/\mathrm{i})\partial/\partial q$. For the parametrized particle, this includes also $\hat{p}_t = (\hbar/\mathrm{i})\partial/\partial t$. Therefore, the quantum version of the constraint (3.8, 3.9) reads

$$\left(\hat{H} - \mathrm{i}\hbar\frac{\partial}{\partial t}\right)\psi(q,t) = 0, \tag{3.14}$$

which is just the Schrödinger equation. Does this mean that t is a dynamical variable in quantum mechanics? The answer is no. We have already mentioned in Section 1.1.2 that time cannot be represented by an operator (e.g. it would be in contradiction with the boundedness of energy). This is the consequence of having an absolute structure in disguise—it remains an absolute structure in quantum theory, in spite of its formal appearance as a quantum variable.

3.1.2 The relativistic particle

We consider a free relativistic particle with mass $m \neq 0$. In units where $c = 1$, its action can be taken to be proportional to the total proper time along its worldline,

$$S = -m\int_{s_1}^{s_2} \mathrm{d}s. \tag{3.15}$$

Using instead of proper time an arbitrary parameter τ for the worldline $x^\mu(\tau)$, the action reads

$$S = -m\int_{\tau_1}^{\tau_2} \mathrm{d}\tau\ \sqrt{-\dot{x}^2}, \tag{3.16}$$

where $\dot{x}^\mu := \mathrm{d}x^\mu/\mathrm{d}\tau$ and $\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu < 0$ (the tangent vector is time-like). One immediately recognizes that the Lagrangian is homogeneous in the velocities and that, therefore, the action is invariant under $\tau \to f(\tau)$. The canonical momenta read

$$p_\mu = \frac{m\dot{x}_\mu}{\sqrt{-\dot{x}^2}}. \tag{3.17}$$

From this expression, it follows immediately that the momenta obey the 'mass-shell condition'

$$p^2 + m^2 = 0. \tag{3.18}$$

In fact, this is a constraint in phase space and thus should be more properly written as $p^2 + m^2 \approx 0$. Because of reparametrization invariance, the canonical Hamiltonian vanishes:

$$H_{\rm c} = p_\mu \dot{x}^\mu - L = \frac{m\dot{x}^2}{\sqrt{-\dot{x}^2}} + m\sqrt{-\dot{x}^2} = 0.$$

In fact,

$$\begin{aligned} H_{\rm c}(x,p) = -p^0\dot{x}^0 + \mathbf{p}\dot{\mathbf{x}} - L &= -p^0\dot{x}^0 + \frac{\mathbf{p}^2\dot{x}^0}{\sqrt{\mathbf{p}^2+m^2}} + m\sqrt{(\dot{x}^0)^2 - \dot{\mathbf{x}}^2} \\ &= \dot{x}^0\left(-p^0 + \sqrt{\mathbf{p}^2+m^2}\right) \approx 0, \end{aligned} \tag{3.19}$$

where the positive square root has been chosen, $p^0 = \sqrt{\mathbf{p}^2+m^2}$, in order to render the energy positive. (The relation (3.17) has been used twice in order to arrive at the expression in the first line.) Analogously to (3.10), one can transform the action into Hamiltonian form,

$$S = \int_{\tau_1}^{\tau_2} \mathrm{d}\tau\ (p_\mu \dot{x}^\mu - NH_{\rm S}), \tag{3.20}$$

where here

$$H_{\rm S} := \eta^{\mu\nu}p_\mu p_\nu + m^2 \approx 0 \tag{3.21}$$

plays the role of the super-Hamiltonian, which is constrained to vanish. The interpretation of the Lagrange multiplier N can be obtained from Hamilton's equations,

$$\dot{x}^0 = \frac{\partial(NH_{\rm S})}{\partial p_0} = -2Np_0,$$

to give

$$N = \frac{\dot{x}^0}{2\sqrt{\mathbf{p}^2+m^2}} = \frac{\dot{x}^0}{2m\gamma} = \frac{1}{2m}\frac{\mathrm{d}s}{\mathrm{d}\tau}, \tag{3.22}$$

where γ is the standard relativistic factor. In contrast to (3.11), the lapse function N here is proportional to the rate of change of proper time (not x^0) with respect to parameter time.[1]

If we apply Dirac's quantization rule to the classical constraint (3.21), we get

$$\hat{H}_{\rm S}\psi(x^\mu) \equiv \left(-\hbar^2\Box + m^2\right)\psi(x^\mu) = 0. \tag{3.23}$$

This is the Klein–Gordon equation for relativistic one-particle quantum mechanics (spinless particles). We emphasize that the classical parameter τ has completely disappeared, since particle trajectories do not exist in quantum theory.

With regard to the Hamiltonian action (3.20), the question arises how x, p, and N must transform under time reparametrizations in order to leave the action invariant;

[1] If we had chosen (3.19) instead of (3.21), we would have found $N = \dot{x}^0$.

cf. Teitelboim (1982). The first-class constraint H_S acts on x and p as follows: one has (neglecting the space–time indices)

$$\delta x(\tau) = \epsilon(\tau)\{x, H_S\} = \epsilon \frac{\partial H_S}{\partial p}, \tag{3.24}$$

$$\delta p(\tau) = \epsilon(\tau)\{p, H_S\} = -\epsilon \frac{\partial H_S}{\partial x}. \tag{3.25}$$

(In the present case this yields $\delta x = 2\epsilon p$, $\delta p = 0$, but we shall keep the formalism general for the moment.) But how does the Lagrange multiplier N transform? We calculate for this purpose

$$\delta S = \int_{\tau_1}^{\tau_2} \mathrm{d}\tau \; (\dot{x}\delta p + p\delta\dot{x} - H_S \delta N - N\delta H_S).$$

The last term is zero, and partial integration of the second term leads to

$$\delta S = \int_{\tau_1}^{\tau_2} \mathrm{d}\tau \; \left(-\epsilon \frac{\partial H_S}{\partial x}\dot{x} - \epsilon \frac{\partial H_S}{\partial p}\dot{p} - H_S \delta N\right) + \left[p\epsilon \frac{\partial H_S}{\partial p}\right]_{\tau_1}^{\tau_2}.$$

In order that only a surface term remains, one has to choose

$$\delta N(\tau) = \dot{\epsilon}(\tau). \tag{3.26}$$

This leads to

$$\delta S = \left[\epsilon(\tau)\left(p\frac{\partial H_S}{\partial p} - H_S\right)\right]_{\tau_1}^{\tau_2}. \tag{3.27}$$

Since the term in brackets gives $p^2 - m^2 \neq 0$, one must demand

$$\epsilon(\tau_1) = 0 = \epsilon(\tau_2)\ ; \tag{3.28}$$

that is, the boundaries must not be transformed.

We note that for a constraint of the form $H_S = \alpha(x)p$, the term in brackets would vanish and there would be no restriction at the boundaries in this case. Constraints of this form arise in electrodynamics and Yang–Mills theories (Gauss constraints) provided the sources are treated dynamically too (otherwise, the constraint would no longer be homogeneous in the momenta; consider, for example, $\nabla \mathbf{E} = \rho$ in electrodynamics).

We shall now show how the gauge can be fixed for the relativistic particle. If a gauge is independent of the lapse function N, it is called a 'canonical gauge', otherwise it is called 'non-canonical'. Consider first a canonical gauge,

$$\chi(x, p, \tau) \approx 0. \tag{3.29}$$

An example would be $x^0 - \tau \approx 0$ (such a gauge was used in the deparametrization of the non-relativistic particle; see the paragraph before (3.12)). A potential problem is that (3.29) holds at all times, including the endpoints, and may thus be in conflict with $\epsilon(\tau_1) = \epsilon(\tau_2) = 0$—since there is no gauge freedom at the endpoints, $\chi \approx 0$ could restrict physically relevant degrees of freedom.

For reparametrization-invariant systems, a canonical gauge must depend explicitly on τ. From the condition that (3.29) be invariant under time evolution,

$$0 \approx \frac{\mathrm{d}\chi}{\mathrm{d}\tau} = \frac{\partial \chi}{\partial \tau} + N\{\chi, H_{\mathrm{S}}\},$$

a τ-independent gauge χ would lead to the unacceptable value $N = 0$ ('freezing' of the motion). (In order for the gauge to break the reparametrization invariance generated by H_{S}, $\{\chi, H_{\mathrm{S}}\}$ must be non-vanishing.) For the relativistic particle, this yields

$$0 \approx \frac{\partial \chi}{\partial \tau} + N\frac{\partial \chi}{\partial x^\mu}\frac{\partial H_{\mathrm{S}}}{\partial p_\mu} = \frac{\partial \chi}{\partial \tau} + 2Np^\mu \frac{\partial \chi}{\partial x^\mu}.$$

For the example $x^0 - \tau \approx 0$, one has $N = 1/2p^0$, in accordance with (3.22).

To avoid potential problems with the boundary, one can look for an equation of second order in ϵ (since there are two conditions, $\epsilon(\tau_1) = 0 = \epsilon(\tau_2)$). As x and p transform proportionally to ϵ, one would have to involve $\ddot{x}$ or $\ddot{p}$, which would render the action functional unnecessarily complicated. Therefore, since $\delta N = \dot{\epsilon}$ (see (3.26)), one can choose the 'non-canonical gauge'

$$\dot{N} = \chi(p, x, N). \tag{3.30}$$

In electrodynamics, A^0 plays the role of N. Therefore, the Lorenz gauge $\partial_\mu A^\mu = 0$ is a non-canonical gauge, whereas the Coulomb gauge $\partial_a A^a = 0$ is an example of a canonical gauge.

Some final remarks are in order; cf. Henneaux and Teitelboim (1992). First, the restriction $\epsilon(\tau_1) = 0 = \epsilon(\tau_2)$ only holds if the action is an integral over the Lagrangian without additional boundary terms. If appropriate boundary terms are present in the action principle, one can relax the condition on ϵ (but to determine these boundary terms, one has to solve first the equations of motion). Second, if such boundary terms are present, one can even choose τ-independent canonical gauges (an extreme choice would be $x^0(\tau) = 0$ for all τ).

3.2 The free bosonic string

Nowadays, superstring theory (or 'M-theory') is considered to be a candidate for a unified theory of all interactions, including quantum gravity. This aspect will be discussed in Chapter 9. In this section, we shall consider the free bosonic string as a model for (canonical) quantum gravity. However, the bosonic string (where no supersymmetry is included) is also used in a heuristic way in the development of superstring theory itself.

In the case of the relativistic particle, the action is proportional to the proper time; see (3.15). A string is a finite one-dimensional object that can be either open (with two ends) or closed (having no ends). A straightforward generalization to the string would thus be to use an action proportional to the area of the worldsheet $\mathcal{M}$, called the Nambu–Goto action,

$$S = -\frac{1}{2\pi\alpha'}\int_{\mathcal{M}} \mathrm{d}^2\sigma\ \sqrt{|\det G_{\alpha\beta}|}. \tag{3.31}$$

Here, $\mathrm{d}^2\sigma := \mathrm{d}\sigma\,\mathrm{d}\tau$ denotes the integration over the parameters of the worldsheet (with both the space part σ and the time part τ chosen to be dimensionless), and $G_{\alpha\beta}$ is the metric on the worldsheet. The string tension is $(2\pi\alpha')^{-1}$; that is, there is a new fundamental parameter α' with dimension length/mass. In the quantum version, the fundamental string length

$$l_{\mathrm{s}} = \sqrt{2\alpha'\hbar} \tag{3.32}$$

will occur. The string propagates in a higher-dimensional space–time, and the worldsheet metric $G_{\alpha\beta}$ $(\alpha, \beta = 1, 2)$ is induced by the metric of the embedding space–time. In the following, we shall assume that the string propagates in D-dimensional Minkowski space, with the worldsheet given by $X^\mu(\sigma,\tau)$, where $\mu = 0, \ldots, D-1$. Denoting the derivative with respect to $\tau \equiv \sigma^0$ by a dot and the derivative with respect to $\sigma \equiv \sigma^1$ by a prime, one has

$$G_{\alpha\beta} = \eta_{\mu\nu}\frac{\partial X^\mu}{\partial\sigma^\alpha}\frac{\partial X^\nu}{\partial\sigma^\beta}, \tag{3.33}$$

$$|\det G_{\alpha\beta}| = -\det G_{\alpha\beta} = (\dot{X}X')^2 - \dot{X}^2(X')^2. \tag{3.34}$$

The embeddings $X^\mu(\sigma,\tau)$ will play the role of the dynamical variables here. The canonical momenta conjugate to them read

$$P_\mu = -\frac{1}{2\pi\alpha'\sqrt{-\det G_{\alpha\beta}}}\left[(\dot{X}X')X'_\mu - (X')^2\dot{X}_\mu\right]. \tag{3.35}$$

From this one gets the conditions

$$P_\mu X^{\mu\prime} = -\frac{1}{2\pi\alpha'\sqrt{-\det G_{\alpha\beta}}}\left[(\dot{X}X')(X')^2 - (X')^2(\dot{X}X')\right] = 0 \tag{3.36}$$

and

$$P_\mu P^\mu = -\frac{(X')^2}{4\pi^2(\alpha')^2}. \tag{3.37}$$

In fact, the last two conditions are just constraints—a consequence of the reparametrization invariance

$$\tau \mapsto \tau'(\tau,\sigma)\ , \quad \sigma \mapsto \sigma'(\tau,\sigma).$$

The constraint (3.37), in particular, is a direct analogue of (3.18).

As expected from the general considerations in Section 3.1.1, the Hamiltonian is constrained to vanish. For the Hamiltonian *density* $\mathcal{H}$, one finds that

$$\mathcal{H} = N\mathcal{H}_\perp + N^1\mathcal{H}_1, \tag{3.38}$$

where N and N^1 are Lagrange multipliers, and

$$\mathcal{H}_\perp = \frac{1}{2}\left(P^2 + \frac{(X')^2}{4\pi^2(\alpha')^2}\right) \approx 0, \tag{3.39}$$

$$\mathcal{H}_1 = P_\mu X^{\mu\prime} \approx 0. \tag{3.40}$$

Quantization of these constraints is formally achieved by imposing the commutation relations

$$[X^\mu(\sigma), P_\nu(\sigma')]|_{\tau=\tau'} = \mathrm{i}\hbar\ \delta^\mu_\nu \delta(\sigma - \sigma') \tag{3.41}$$

and implementing the constraints à la Dirac as restrictions on physically allowed wave functionals,

$$\hat{\mathcal{H}}_\perp \Psi[X^\mu(\sigma)] \equiv \frac{1}{2}\left(-\hbar^2 \frac{\delta^2\Psi}{\delta X^2} + \frac{(X')^2\Psi}{4\pi^2(\alpha')^2}\right) = 0, \tag{3.42}$$

and

$$\hat{\mathcal{H}}_1 \Psi[X^\mu(\sigma)] \equiv \frac{\hbar}{\mathrm{i}} X^{\mu\prime} \frac{\delta\Psi}{\delta X^\mu} = 0, \tag{3.43}$$

where the factor ordering has been chosen such that the momenta are on the right. Note that in contrast to the examples in Section 3.1, one has now to deal with *functional* derivatives, defined by the Taylor expansion

$$\Psi[\phi(\sigma) + \eta(\sigma)] = \Psi[\phi(\sigma)] + \int \mathrm{d}\sigma\ \frac{\delta\Psi}{\delta\phi(\sigma)}\eta(\sigma) + \cdots . \tag{3.44}$$

The above implementation of the constraints is only possible if there are no *anomalies*; see the end of this section and Section 5.3.5. An important property of the quantized string is that such anomalies in fact occur, *preventing* the validity of all quantum equations (3.42) and (3.43). Equations such as (3.42) and (3.43) will occur in several places later and will be discussed further there, for example in the context of parametrized field theories (Section 3.3).

Note that this level of quantization corresponds to a 'first-quantized string', in analogy to first quantization of point particles (Section 3.1). The usual 'second quantization' would mean elevating the wave functions $\Psi[X^\mu(\sigma)]$ themselves into operators ('string field theory'; see Section 9.3). It must also be emphasized that 'first' and 'second' are at best heuristic notions since there is just one quantum theory (cf. in this context Zeh (2003)).

In the following, we shall briefly discuss the connection with the standard textbook treatment of the bosonic string; see, for example, Polchinski (1998*a*). This will also be a useful preparation for the discussion of string theory in Chapter 9. Usually one starts with the Polyakov action for the bosonic string,

$$S_\mathrm{P} = -\frac{1}{4\pi\alpha'}\int_\mathcal{M} \mathrm{d}^2\sigma\ \sqrt{h} h^{\alpha\beta}(\sigma,\tau)\partial_\alpha X^\mu \partial_\beta X_\mu, \tag{3.45}$$

where $h_{\alpha\beta}$ denotes the *intrinsic* (not induced) metric on the worldsheet, and $h := |\det h_{\alpha\beta}|$. In contrast to the induced metric, the intrinsic metric consists of independent degrees of freedom with respect to which the action can be varied. The action (3.45) can be interpreted as describing 'two-dimensional gravity coupled to D massless scalar fields'. Since the Einstein–Hilbert action is a topological invariant in two dimensions, there is no pure gravity term present, and only the coupling of the metric to the X^μ remains in (3.45). One can also take into account a cosmological term; see Chapter 9. In contrast to (3.31), the Polyakov action is much easier to handle, especially when used in a path integral.

The Polyakov action has many invariances. First, it is invariant with respect to diffeomorphisms on the worldsheet. Second, and most importantly, it possesses *Weyl invariance*, that is, an invariance under the transformations

$$h_{\alpha\beta}(\sigma,\tau) \mapsto \mathrm{e}^{2\omega(\sigma,\tau)} h_{\alpha\beta}(\sigma,\tau) \tag{3.46}$$

with an arbitrary function $\omega(\sigma,\tau)$. This is a special feature of two dimensions, where $\sqrt{h}h^{\alpha\beta} \mapsto \sqrt{h}h^{\alpha\beta}$. It means that distances on the worldsheet have no intrinsic physical meaning. In addition, there is the Poincaré symmetry of the background Minkowski space–time, which is of minor interest here.

Defining the two-dimensional energy–momentum tensor according to[2]

$$T_{\alpha\beta} = -\frac{4\pi\alpha'}{\sqrt{h}}\frac{\delta S_{\mathrm{P}}}{\delta h^{\alpha\beta}} = \partial_\alpha X^\mu \partial_\beta X_\mu - \frac{1}{2} h_{\alpha\beta} h^{\gamma\delta} \partial_\gamma X^\mu \partial_\delta X_\mu, \tag{3.47}$$

one finds

$$h^{\alpha\beta} T_{\alpha\beta} = 0. \tag{3.48}$$

This tracelessness of the energy–momentum tensor is a consequence of Weyl invariance. This can be easily seen as follows. Consider a general variation of the action,

$$\delta S = \int \mathrm{d}^2\sigma \; \frac{\delta S}{\delta h^{\alpha\beta}} \delta h^{\alpha\beta} \propto \int \mathrm{d}^2\sigma \; \sqrt{h} T_{\alpha\beta} \delta h^{\alpha\beta}.$$

Under (3.46), we have

$$\delta h^{\alpha\beta} = -2(\delta\omega) h^{\alpha\beta}.$$

Therefore, the demand that $\delta S = 0$ under (3.46) leads to (3.48). In the quantum theory, a 'Weyl anomaly' may occur, in which the trace of the energy–momentum tensor is proportional to $\hbar$ times the two-dimensional Ricci scalar; see (2.125). The demand for this anomaly to vanish leads to restrictions on the parameters of the theory; see below.

Using the field equations $\delta S_{\mathrm{P}}/\delta h_{\alpha\beta} = 0$, one finds

$$0 = T_{\alpha\beta}. \tag{3.49}$$

In a sense, these are the Einstein equations with the left-hand side missing, since the Einstein–Hilbert action is a topological invariant. As (3.49) has no second time derivatives, it is in fact a constraint—a consequence of diffeomorphism invariance. From (3.49), one can easily derive that

$$\det G_{\alpha\beta} = \frac{h}{4}(h^{\alpha\beta} G_{\alpha\beta})^2, \tag{3.50}$$

where $G_{\alpha\beta}$ is the induced metric (3.33). Inserting this into (3.45) gives back the action (3.31). Therefore, 'on-shell' (i.e., using the classical equations), the two actions are equivalent.

[2]Compared with GR, there is an additional factor -2π here, which is introduced for convenience.

The constraints (3.36) and (3.37) can also be found directly from (3.45)—defining the momenta conjugate to X^μ in the usual manner—after use has been made of (3.49). One can thus formulate instead of (3.45) an alternative canonical action principle,

$$S = \int_{\mathcal{M}} \mathrm{d}^2\sigma \, (P_\mu \dot{X}^\mu - N\mathcal{H}_\perp - N^1\mathcal{H}_1), \tag{3.51}$$

where $\mathcal{H}_\perp$ and $\mathcal{H}_1$ are given by (3.39) and (3.40), respectively.

In the standard treatment of the bosonic string, the 'gauge freedom' (with respect to two diffeomorphisms and one Weyl transformation) is fixed by the choice $h_{\alpha\beta} = \eta_{\alpha\beta} = \mathrm{diag}(-1, 1)$. Instead of (3.45) one then has

$$S_{\mathrm{P}} = -\frac{1}{4\pi\alpha'} \int \mathrm{d}^2\sigma \, \eta^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X_\mu, \tag{3.52}$$

that is, the action for D free scalar fields in two-dimensional Minkowski space. In two dimensions, there is a remaining symmetry which leaves the gauge-fixed action invariant—the *conformal transformations* (Kastrup 2008). These are angle-preserving coordinate transformations[3] which change the metric by a factor $\mathrm{e}^{2\omega(\sigma,\tau)}$; they can therefore be compensated by a Weyl transformation, and the action (3.52) is invariant under this combined transformation. A field theory with this invariance is called a 'conformal field theory' (CFT). A particular feature of two dimensions is that the conformal group is infinite-dimensional, giving rise to infinitely many conserved charges (see below).

Consider in the following the case of open strings where $\sigma \in (0, \pi)$. (Closed strings lead to a doubling of the degrees of freedom corresponding to left- and right-movers.) The Hamiltonian of the gauge-fixed theory reads

$$H = \frac{1}{4\pi\alpha'} \int_0^\pi \mathrm{d}\sigma \left(\dot{X}^2 + (X')^2 \right). \tag{3.53}$$

Introducing the components of the energy–momentum tensor with respect to the light-cone coordinates $\sigma^- = \tau - \sigma$ and $\sigma^+ = \tau + \sigma$, it is convenient to define the quantities ($m \in \mathbb{Z}$)

$$L_m = \frac{1}{2\pi\alpha'} \int_0^\pi \mathrm{d}\sigma \ \left(\mathrm{e}^{\mathrm{i}m\sigma} T_{++} + \mathrm{e}^{-\mathrm{i}m\sigma} T_{--} \right). \tag{3.54}$$

One recognizes that $L_0 = H$. Because the energy–momentum tensor vanishes as a constraint, this holds also for the L_m, that is, $L_m \approx 0$. The L_m obey the classical *Virasoro algebra*

$$\{L_m, L_n\} = -\mathrm{i}(m - n)L_{m+n}, \tag{3.55}$$

exhibiting the fact that they generate the group of conformal transformations (the residual symmetry of the gauge-fixed action). The $\{L_n\}$ are the infinitely many conserved charges mentioned above. It turns out that quantization does not preserve this

[3] In GR, the term 'conformal transformation' is usually employed for what is called a Weyl transformation here.

algebra but yields an additional term called the 'anomaly', 'central term', or 'Schwinger term',

$$[\hat{L}_m, \hat{L}_n] = (m-n)\hbar\hat{L}_{m+n} + \frac{c\hbar^2}{12}(m^3 - m)\delta_{m+n,0}, \tag{3.56}$$

where c is the *central charge*. For the case of the free fields X^μ, it is equal to the number of space–time dimensions, $c = D$. Due to the presence of this extra term, one cannot implement the constraints $L_m \approx 0$ in the quantum theory as restrictions on wave functions; that is, one cannot have $\hat{L}_m|\psi\rangle = 0$ for all m. Instead, one can choose

$$\hat{L}_n|\psi\rangle = 0,\ n > 0\ , \quad \hat{L}_0|\psi\rangle = a\hbar|\psi\rangle. \tag{3.57}$$

It turns out that Weyl invariance can only be preserved at the quantum level for $a = 1$ and $D = 26$; see, for example, Polchinski (1998*a*) and Chapter 9. This is achieved by the presence of Faddeev–Popov ghost degrees of freedom whose central charge cancels against the central charge c of the fields X^μ precisely for $D = 26$. It is most elegantly treated by 'BRST quantization' (see Section 9.2), leading to an equation of the form $Q_{\mathrm{B}}|\Psi_{\mathrm{tot}}\rangle = 0$, where Q_{B} is the BRST charge. This weaker condition replaces the direct quantum implementation of the constraints.

Going back to the classical theory, one can also define the quantities

$$\tilde{L}_n = \frac{1}{2}\int \mathrm{d}\sigma\ \mathrm{e}^{\mathrm{i}n\sigma}\left(\sqrt{(X')^2}\mathcal{H}_\perp + \mathcal{H}_1\right). \tag{3.58}$$

Using the Poisson-bracket relations between the constraints $\mathcal{H}_\perp$ and $\mathcal{H}_1$ (see in particular Section 3.3), one can show that

$$\{\tilde{L}_m, \tilde{L}_n\} = -\mathrm{i}(m-n)\tilde{L}_{m+n}, \tag{3.59}$$

which coincides with the Virasoro algebra (3.55). In fact, for the gauge fixing considered here—leading to (3.52)—one has $\tilde{L}_n = L_n$. The result (3.56) then shows that the naive implementation of the constraints (3.42) and (3.43) may be inconsistent. This is a general problem in the quantization of constrained systems and will be discussed further in Section 5.3. Kuchař and Torre (1991) have treated the bosonic string as a model for quantum gravity. They showed that a covariant (covariant with respect to diffeomorphisms of the worldsheet) quantization is possible; that is, there exists a quantization procedure in which the algebra of constraints contains no anomalous terms. This is achieved by extracting internal time variables ('embeddings') which are not represented as operators.[4] A potential problem is the dependence of the theory on the choice of embedding. This is in fact a general problem; see Section 5.2. Kuchař and Torre make use of the fact that string theory is an 'already-parametrized theory', which brings us to a detailed discussion of parametrized field theories in the next section.

[4]The anomaly is still present in a subgroup of the conformal group, but it does not disturb the Dirac quantization of the constraints.

3.3 Parametrized field theories

This example is a generalization of the parametrized non-relativistic particle discussed in Section 3.1.1. As it will be field-theoretic by nature, it has similarities with the bosonic string discussed in the last section, but with notable differences. General references to parametrized field theories are provided by Kuchař (1973, 1981), from which the following material is partially drawn.

The starting point is a real scalar field in Minkowski space, $\phi(X^\mu)$, where the standard inertial coordinates are called $X^\mu \equiv (T, X^a)$ here. We now introduce arbitrary (in general curved) coordinates $x^\mu \equiv (t, x^a)$ and let the X^μ depend parametrically on x^μ. This is analogous to the dependence $t(\tau)$ in Section 3.1.1. The functions $X^\mu(x^\nu)$ describe a family of hypersurfaces in Minkowski space parametrized by $x^0 \equiv t$ (we shall restrict ourselves to the space-like case). Analogously to Section 3.1.1, the standard action for a scalar field is rewritten in terms of the arbitrary coordinates x^μ. This yields

$$S = \int \mathrm{d}^4 X \; \mathcal{L}\left(\phi, \frac{\partial \phi}{\partial X^\mu}\right) =: \int \mathrm{d}^4 x \; \tilde{\mathcal{L}}, \tag{3.60}$$

where

$$\tilde{\mathcal{L}}(\phi, \phi_{,a}, \dot{\phi}; X^\mu_{,a}, \dot{X}^\mu) = J\mathcal{L}\left(\phi, \phi_{,\nu}\frac{\partial x^\nu}{\partial X^\mu}\right), \tag{3.61}$$

and J denotes the Jacobi determinant of the X with respect to the x (a dot denotes a differentiation with respect to t, and a comma denotes a differentiation with respect to the x^ν). Instead of calculating directly the momentum canonically conjugate to X^μ, it is more appropriate to consider first the Hamiltonian density $\tilde{\mathcal{H}}$ corresponding to $\tilde{\mathcal{L}}$ with respect to ϕ,

$$\begin{aligned}
\tilde{\mathcal{H}} &= \tilde{p_\phi}\dot{\phi} - \tilde{\mathcal{L}} = J\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\dot{\phi} - J\mathcal{L} \\
&= J\frac{\partial x^0}{\partial X^\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial\phi/\partial X^\mu\right)}\frac{\partial \phi}{\partial X^\nu} - \delta^\mu_\nu \mathcal{L}\right)\dot{X}^\nu \\
&\equiv J\frac{\partial x^0}{\partial X^\mu}T^\mu{}_\nu \dot{X}^\nu.
\end{aligned} \tag{3.62}$$

Both J and the canonical energy–momentum tensor $T^\mu{}_\nu$ do not, in fact, depend on the 'kinematical velocities' $\dot{X}^\mu$. This can be seen as follows. The Jacobi determinant J can be written as

$$J = \epsilon_{\rho\nu\lambda\sigma}\frac{\partial X^\rho}{\partial x^0}\frac{\partial X^\nu}{\partial x^1}\frac{\partial X^\lambda}{\partial x^2}\frac{\partial X^\sigma}{\partial x^3},$$

from which one gets

$$J\frac{\partial x^0}{\partial X^\mu} = \epsilon_{\mu\nu\lambda\sigma}\frac{\partial X^\nu}{\partial x^1}\frac{\partial X^\lambda}{\partial x^2}\frac{\partial X^\sigma}{\partial x^3},$$

which is just the vectorial surface element on $t =$ constant, which does not depend on the $\dot{X}^\mu$. For the same reason, the energy–momentum tensor does not depend on these velocities.

As a generalization of (3.7) and (3.9), one may introduce the kinematical momenta Π_ν via the constraint

$$\mathcal{H}_\nu := \Pi_\nu + J\frac{\partial x^0}{\partial X^\mu} T^\mu{}_\nu \approx 0. \tag{3.63}$$

Taking then the action

$$S = \int \mathrm{d}^4x\ (\tilde{p}_\phi \dot{\phi} - \tilde{\mathcal{H}}),$$

inserting (3.62) for $\tilde{\mathcal{H}}$, and adding the constraints (3.63) with Lagrange multipliers N^ν, one gets the action principle

$$S = \int \mathrm{d}^4x\ (\tilde{p}_\phi \dot{\phi} + \Pi_\nu \dot{X}^\nu - N^\nu \mathcal{H}_\nu). \tag{3.64}$$

This is the result that one would also get by defining the kinematical momenta directly from (3.61). It is analogous to (3.10).

It is convenient to decompose (3.63) into components orthogonal and parallel to the hypersurfaces $x^0 =$ constant. Introducing the normal vector n^μ (with $\eta_{\mu\nu} n^\mu n^\nu = -1$) and the tangential vectors $X^\nu_{,a}$ (obeying $n_\nu X^\nu_{,a} = 0$), one gets the constraints

$$\mathcal{H}_\perp := \mathcal{H}_\nu n^\nu \approx 0, \tag{3.65}$$

$$\mathcal{H}_a := \mathcal{H}_\nu X^\nu_{,a} \approx 0. \tag{3.66}$$

Equations (3.65) and (3.66) are called the Hamiltonian constraint and the momentum constraints, respectively. The latter are also called diffeomorphism constraints, because they can generate infinitesimal diffeomorphisms; see (3.82) below. The whole set of constraints is fully analogous to the corresponding constraints (3.39) and (3.40) in string theory.

The action (3.64) then reads

$$S = \int \mathrm{d}^4x\ (\tilde{p}_\phi \dot{\phi} + \Pi_\nu \dot{X}^\nu - N\mathcal{H}_\perp - N^a \mathcal{H}_a). \tag{3.67}$$

To interpret the Lagrange multipliers N and N^a, we vary this action with respect to Π_ν and obtain

$$\dot{X}^\nu =: t^\nu = N n^\nu + N^a X^\nu_{,a}. \tag{3.68}$$

The geometric interpretation is depicted in Fig. 3.1. $\dot{X}^\nu$ is a vector that points from a point with (spatial) coordinates x^a on $t =$ constant to a point with the *same* coordinates on a neighbouring hypersurface $t + \mathrm{d}t =$ constant. The purely temporal distance between the hypersurfaces is given by $N\mathrm{d}t$, and therefore N is called the *lapse function*. Similarly, N^a is a vector that points from the point with coordinates x^a on $t =$ constant to the point on the same hypersurface from which the normal is erected to reach the point with the same coordinates x^a on $t + \mathrm{d}t =$ constant. It is called the *shift vector*.

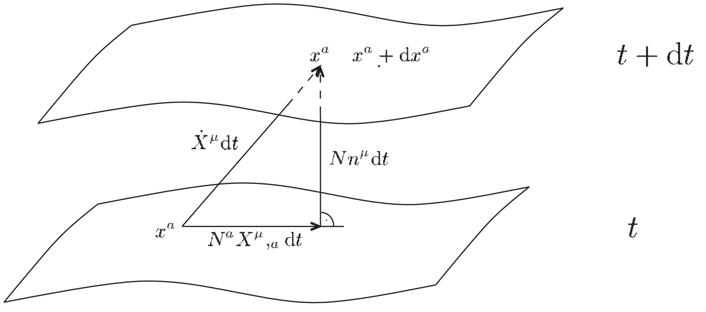

Fig. 3.1 Geometric interpretation of lapse and shift.

Instead of Minkowski space, one can also choose an arbitrary curved background space–time with metric $g_{\mu\nu}$ for the embedding. Denoting the spatial metric induced on the hypersurface by h_{ab}, that is

$$h_{ab} = g_{\mu\nu} \frac{\partial X^\mu}{\partial x^a} \frac{\partial X^\nu}{\partial x^b}, \tag{3.69}$$

the four-dimensional line element can be decomposed as follows:

$$\begin{aligned} \mathrm{d}s^2 &= g_{\mu\nu}\, \mathrm{d}x^\mu\, \mathrm{d}x^\nu = -N^2\, \mathrm{d}t^2 + h_{ab}(\mathrm{d}x^a + N^a\, \mathrm{d}t)(\mathrm{d}x^b + N^b\, \mathrm{d}t) \\ &= (h_{ab}N^aN^b - N^2)\mathrm{d}t^2 + 2h_{ab}N^a\, \mathrm{d}x^b\, \mathrm{d}t + h_{ab}\, \mathrm{d}x^a\, \mathrm{d}x^b. \end{aligned} \tag{3.70}$$

The action (3.67) is invariant under the reparametrizations

$$\begin{aligned} x^0 &\to x^{0'} = x^0 + f(x^a), \\ x^a &\to x^{a'} = g(x^b) \end{aligned} \tag{3.71}$$

with arbitrary functions (obeying standard differentiability conditions) f and g. This is not equivalent to the full set of space–time diffeomorphisms; see the discussion at the end of this section.

A simple example of the above procedure is the case of a massless scalar field on (1+1)-dimensional Minkowski space–time (Kuchař 1973, 1981). Its Lagrangian reads

$$\mathcal{L}\left(\phi, \frac{\partial\phi}{\partial T}, \frac{\partial\phi}{\partial X}\right) = -\frac{1}{2}\eta^{\mu\nu}\frac{\partial\phi}{\partial X^\mu}\frac{\partial\phi}{\partial X^\nu} = \frac{1}{2}\left[\left(\frac{\partial\phi}{\partial T}\right)^2 - \left(\frac{\partial\phi}{\partial X}\right)^2\right]. \tag{3.72}$$

For the Jacobi determinant, we have

$$J = \dot{T}X' - T'\dot{X},$$

and for the components of the normal vector,

$$n^T = \frac{X'}{\sqrt{X'^2 - T'^2}}\,, \quad n^X = \frac{T'}{\sqrt{X'^2 - T'^2}}.$$

(Dots denote derivatives with respect to $x^0 \equiv t$, and primes denote derivatives with respect to $x^1 \equiv x$.) The energy–momentum tensor assumes the well-known form for a free scalar field. If the above procedure is followed, one finds for the constraints (3.65) and (3.66) the expressions

$$\mathcal{H}_\perp = \frac{1}{\sqrt{X'^2 - T'^2}} \left(X'\Pi_T + T'\Pi_X + \frac{1}{2}(\tilde{p}_\phi^2 + \phi'^2) \right), \tag{3.73}$$

$$\mathcal{H}_1 = T'\Pi_T + X'\Pi_X + \phi'\tilde{p}_\phi. \tag{3.74}$$

(The space-like nature of the hypersurfaces guarantees that the radicand of the square root is positive.) One recognizes from (3.73) that the kinematical momenta Π_T and Π_X enter the Hamiltonian constraint *linearly*. This is different from the string case (3.39) and distinguishes, in fact, a parametrized theory from a theory which is *intrinsically* reparametrization invariant, that is, a theory which is background independent.

Quantization is performed by imposing the formal commutators

$$[X^\mu(\mathbf{x}), \Pi_\nu(\mathbf{y})] = \mathrm{i}\hbar\ \delta^\mu_\nu \delta(\mathbf{x} - \mathbf{y}) \tag{3.75}$$

and

$$[\phi(\mathbf{x}), \tilde{p}_\phi(\mathbf{y})] = \mathrm{i}\hbar\ \delta(\mathbf{x} - \mathbf{y}). \tag{3.76}$$

From (3.65) and (3.66), one then finds, according to Dirac's prescription, the quantum constraints

$$\mathcal{H}_\perp \Psi[\phi(x), X^\mu(x)] = 0, \tag{3.77}$$

$$\mathcal{H}_a \Psi[\phi(x), X^\mu(x)] = 0. \tag{3.78}$$

In the above example of the free scalar field, the constraints read

$$\begin{aligned}\mathcal{H}_\perp \Psi = {} & \frac{1}{\sqrt{X'^2 - T'^2}} \left(-\mathrm{i}\hbar X'(x) \frac{\delta}{\delta T(x)} - \mathrm{i}\hbar T'(x) \frac{\delta}{\delta X(x)} \right. \\ & \left. + \frac{1}{2} \left[-\hbar^2 \frac{\delta^2}{\delta\phi(x)\delta\phi(x)} + \phi'^2(x) \right] \right) \Psi[\phi(x), T(x), X(x)] = 0, \end{aligned} \tag{3.79}$$

$$\mathcal{H}_1 \Psi = \frac{\hbar}{\mathrm{i}} \left(T'(x) \frac{\delta}{\delta T(x)} + X'(x) \frac{\delta}{\delta X(x)} + \phi'(x) \frac{\delta}{\delta \phi(x)} \right) \Psi = 0. \tag{3.80}$$

The functional derivatives occurring in these equations are treated formally, that is, as if they were ordinary derivatives, and all derivatives are put to the right. One should be aware that there is always the problem of factor ordering (as in quantum mechanics) and that singularities arise if a functional derivative is taken of an ordinary function with respect to the same argument. This is a general problem of functional differential equations and will be discussed further in Section 5.3.

The above equations are very different from the equations that one would get from the standard Lagrangian (3.72) or the corresponding action. The reason is that the wave functional is usually evolved along *flat* hypersurfaces T = constant only, whereas in the parametrized version it can be evolved along *any* family of space-like hypersurfaces. The latter description needs two functions $X(x)$ and $T(x)$, also called a 'many-fingered time' or 'bubble-time' description. As with the non-relativistic particle, the parametrized theory can easily be deparametrized. Choosing $x = X$

as a coordinate on the hypersurfaces and evolving the wave functionals along flat hypersurfaces described by $T(x) = T_0 \in (-\infty, \infty)$, one finds (Kuchař 1973)

$$\mathrm{i}\hbar \frac{\partial \Psi}{\partial T_0} = \frac{1}{2} \int \mathrm{d}X \left(-\hbar^2 \frac{\delta^2}{\delta \phi^2(X)} + \left[\frac{\partial \phi}{\partial X} \right]^2 \right) \Psi, \tag{3.81}$$

which is just the ordinary (functional) Schrödinger equation for the massless scalar field. Note that this is only one equation instead of the infinitely many equations (3.79) and (3.80).

The general interpretation of the momentum constraint can be easily recognized from the example (3.80). Performing an infinitesimal coordinate transformation on $T =$ constant, $x \to \bar{x} = x + \delta N^1(x)$, one gets

$$T(x) \to T(x + \delta N^1(x)) = T(x) + T'(x)\delta N^1(x)$$

and similar equations for $X(x)$ and $\phi(x)$. For the wave functional, the transformation yields

$$\begin{aligned} \Psi \to \Psi[T(x) + T'(x)\delta N^1(x), \ldots] &= \Psi[T(x), \ldots] \\ &+ \int \mathrm{d}x \, \left(T'(x) \frac{\delta \Psi}{\delta T(x)} + \cdots \right) \delta N^1(x). \end{aligned} \tag{3.82}$$

Therefore, the momentum constraint (3.80) enforces the independence of Ψ under infinitesimal coordinate transformations on the hypersurfaces, for then the integrand in this equation vanishes and the wave functional remains unchanged. The constraint acts as a generator of infinitesimal diffeomorphisms.

Going back to the general action (3.67), one finds that the Hamiltonian is, as expected, a linear combination of constraints,

$$H = \int \mathrm{d}^3 x \, \left(N \mathcal{H}_\perp + N^a \mathcal{H}_a \right). \tag{3.83}$$

Dynamical consistency of a constrained system is only obtained if the constraints are preserved in time (here, with respect to the time parameter x^0). This is the case only if the Poisson brackets between all constraints are combinations of the constraints themselves. One finds in fact the Poisson-bracket algebra (Dirac 1964)

$$\{\mathcal{H}_\perp(\mathbf{x}), \mathcal{H}_\perp(\mathbf{y})\} = -\sigma \delta_{,a}(\mathbf{x}, \mathbf{y}) \left(h^{ab}(\mathbf{x}) \mathcal{H}_b(\mathbf{x}) + h^{ab}(\mathbf{y}) \mathcal{H}_b(\mathbf{y}) \right), \tag{3.84}$$

$$\{\mathcal{H}_a(\mathbf{x}), \mathcal{H}_\perp(\mathbf{y})\} = \mathcal{H}_\perp(\mathbf{x}) \delta_{,a}(\mathbf{x}, \mathbf{y}), \tag{3.85}$$

$$\{\mathcal{H}_a(\mathbf{x}), \mathcal{H}_b(\mathbf{y})\} = \mathcal{H}_b(\mathbf{x}) \delta_{,a}(\mathbf{x}, \mathbf{y}) + \mathcal{H}_a(\mathbf{y}) \delta_{,b}(\mathbf{x}, \mathbf{y}), \tag{3.86}$$

with the derivatives all acting on $\mathbf{x}$. We have introduced here a space–time metric with signature diag$(\sigma, 1, 1, 1)$ in order to exhibit the difference between the Lorentzian case ($\sigma = -1$, the relevant case here) and the Euclidean case ($\sigma = 1$). This algebra will play a crucial role in canonical gravity; see Chapters 4–6. It is often convenient to work with a 'smeared' version of the constraints, that is,

$$\mathcal{H}[N] = \int \mathrm{d}^3x\ N(\mathbf{x})\mathcal{H}_\perp(\mathbf{x}) \ , \quad \mathcal{H}[N^a] = \int \mathrm{d}^3x\ N^a(\mathbf{x})\mathcal{H}_a(\mathbf{x}). \tag{3.87}$$

The constraint algebra then reads

$$\{\mathcal{H}[N], \mathcal{H}[M]\} = \mathcal{H}[K^a] \ , \quad K^a = -\sigma h^{ab}(NM_{,b} - N_{,b}M), \tag{3.88}$$

$$\{\mathcal{H}[N^a], \mathcal{H}[N]\} = \mathcal{H}[M] \ , \quad M = N^a N_{,a} \equiv \mathcal{L}_{\mathbf{N}} N, \tag{3.89}$$

$$\{\mathcal{H}[N^a], \mathcal{H}[M^b]\} = \mathcal{H}[\mathbf{K}] \ , \quad \mathbf{K} = [\mathbf{N}, \mathbf{M}] \equiv \mathcal{L}_{\mathbf{N}}\mathbf{M}, \tag{3.90}$$

where $\mathcal{L}$ denotes here the Lie derivative. Some remarks are in order:

1. Since the algebra closes, the constraints are all of first class. (See Section 3.5 for a brief review of constrained systems.)
2. The algebra is *not* a Lie algebra, since (3.88) contains the (inverse) three-metric $h^{ab}(\mathbf{x})$ of the hypersurfaces x^0 = constant (i.e. one has structure functions depending on the canonical variables instead of structure constants). An exception is two-dimensional space–time (Teitelboim 1984).
3. The signature σ of the embedding space–time can be read off directly from (3.88).
4. The subalgebra of the diffeomorphism constraints *is* a Lie algebra, as can be seen from (3.90). Equation (3.89) means that the flow of the Hamiltonian constraint does not leave the constraint hypersurface of the diffeomorphism constraints invariant. Moreover, this equation expresses the fact that $\mathcal{H}_\perp$ transforms under diffeomorphisms as a scalar density of weight one; this follows from (3.85) after integration with respect to the shift vector,

$$\delta\mathcal{H}_\perp(\mathbf{y}) = \int \mathrm{d}^3x\ N^a(\mathbf{x})\{\mathcal{H}_\perp(\mathbf{y}), \mathcal{H}_a(\mathbf{x})\} = \partial_a(N^a\mathcal{H}_\perp)(\mathbf{y}).$$

5. The algebra is the same as for the corresponding constraints in the case of the bosonic string; that is, it is in two dimensions equivalent to the Virasoro algebra (3.55). The reason is its general geometric interpretation, to be discussed in the following.

It turns out that the above algebra has a purely kinematical interpretation. It is just the algebra of surface deformations for hypersurfaces which are embedded in a Riemannian (or pseudo-Riemannian) space. If a hypersurface is again described by $X^\mu(\mathbf{x})$, the generators of coordinate transformations *on* the hypersurface are given by

$$X_{ax} := X^\mu_{,a}(\mathbf{x})\frac{\delta}{\delta X^\mu(\mathbf{x})},$$

while the generators of the normal deformations are given by

$$X_x := n^\mu(\mathbf{x})\frac{\delta}{\delta X^\mu(\mathbf{x})},$$

with the normal vector obeying the normalization condition $n^\mu n_\mu = \sigma$. Kuchař (1973) calculated the algebra,

$$[X_x, X_y] = \sigma\delta_{,a}(\mathbf{x}, \mathbf{y})\left(h^{ab}(\mathbf{x})X_{bx} + h^{ab}(\mathbf{y})X_{by}\right), \tag{3.91}$$

$$[X_{ax}, X_y] = -X_x\delta_{,a}(\mathbf{x}, \mathbf{y}), \tag{3.92}$$

$$[X_{ax}, X_{by}] = -X_{bx}\delta_{,a}(\mathbf{x}, \mathbf{y}) - X_{ay}\delta_{,b}(\mathbf{x}, \mathbf{y}). \tag{3.93}$$

Up to a sign, this algebra has the same structure as the constraint algebra (3.84)–(3.86). The reason for the different sign is the relation

$$[X_f, X_g] = -X_{\{f,g\}},$$

where $\{f, g\}$ denotes the Poisson bracket between f and g. The constraints $\mathcal{H}_\perp$ and $\mathcal{H}_a$ generate, in fact, the algebra of hypersurface deformations, which are given by (3.71) and which are *not* identical to the algebra of space–time diffeomorphisms; cf. also Chapter 4. That surface deformations form a larger class of transformations than space–time diffeomorphisms can be seen if one considers surfaces which intersect each other at a point P. Under surface deformations this point is shifted to two *different* points, depending on which of the two surfaces is deformed first. On the other hand, a space–time diffeomorphism shifts each point in a unique way independent of the surface on which it lies; space–time diffeomorphisms thus induce only special surface deformations.

In Section 4.1, we will show that one can construct the theory of GR from the above constraint algebra provided that the three-metric h_{ab} and its canonical momentum are the only canonical variables. Before this is done, we shall discuss in the next section a 'relational' mechanical model that exhibits some interesting features relevant to the quantization of gravity.

3.4 Relational dynamical systems

Newtonian mechanics needs for its formulation the concepts of absolute space and absolute time. This was criticized by some of Newton's contemporaries, notably Leibniz, who insisted that only observable quantities should appear in the fundamental equations. In the nineteenth century, Ernst Mach emphasized that the concepts of absolute space and time should be abandoned altogether and that physics should only use relational concepts.

Let us consider, in a gedanken experiment, two successive 'snapshots' of the universe separated by a short time interval (Barbour 1986). The universe is considered for simplicity as a collection of n particles with masses m_i, $i = 1, \ldots, n$, evolving in Euclidean space under the influence of Newtonian gravity. Due to the short time interval, the relative distances will be only slightly different. Can one predict the future evolution from these two observations? The definite answer is *no*, because the two sets of relative separations give no information about the angular momentum or kinetic energy of the system, both of which affect the future evolution.

This 'Poincaré defect' (because Poincaré identified this lack of predictability; see Poincaré (2009)) motivated Barbour and Bertotti (1982) to look for a modification of Newtonian mechanics in which the future can be predicted solely on the basis of *relative* separations (and their rates of change). The key idea is to introduce a 'gauge freedom' with respect to translations and rotations (because these transformations leave the relative distances invariant) and the choice of the time parameter τ. The theory should thus be invariant under the following gauge transformations:

$$\mathbf{x}_k \mapsto \mathbf{x}'_k = \mathbf{x}_k + \mathbf{a}(\tau) + \boldsymbol{\alpha}(\tau) \times \mathbf{x}_k, \tag{3.94}$$

where $\mathbf{a}$ parametrizes translations, $\boldsymbol{\alpha}$ parametrizes rotations, and $\mathbf{x}_k$ is the position vector of particle k. They depend on the 'label time' τ, which can be arbitrarily reparametrized:

$$\tau \mapsto f(\tau)\ , \quad \dot{f} > 0. \tag{3.95}$$

Due to (3.94), one has only $3n-6$ parameters to describe the relative distances instead of the original $3n$. Equations (3.94) and (3.95) define the 'Leibniz group' (Barbour and Bertotti 1982, Barbour 1986). One can now define a total velocity for each particle according to

$$\frac{D\mathbf{x}_k}{D\tau} := \frac{\partial \mathbf{x}_k}{\partial \tau} + \dot{\mathbf{a}}(\tau) + \dot{\boldsymbol{\alpha}}(\tau) \times \mathbf{x}_k, \tag{3.96}$$

in which the first term on the right-hand side denotes the rate of change in some chosen frame, and the second and third terms the rate of change due to a τ-dependent change of frame. This velocity is not yet gauge invariant. A gauge-invariant quantity can be constructed by minimizing the 'kinetic energy'

$$\sum_{k=1}^{n} \frac{D\mathbf{x}_k}{D\tau}\frac{D\mathbf{x}_k}{D\tau}$$

with respect to $\mathbf{a}$ and $\boldsymbol{\alpha}$. This procedure is also called 'horizontal stacking' (Barbour 1986). Intuitively, it can be understood as putting two slides with the particle positions marked on them on top of each other and moving them relative to each other until the centres of mass coincide and there is no overall rotation. The result of the horizontal stacking is a gauge-invariant 'intrinsic velocity', $\mathrm{d}\mathbf{x}/\mathrm{d}\tau$. Having these velocities for each particle at one's disposal, one can construct the kinetic term

$$T = \frac{1}{2}\sum_{k=1}^{n} m_k \left(\frac{\mathrm{d}\mathbf{x}_k}{\mathrm{d}\tau}\right)^2 . \tag{3.97}$$

The potential is the standard Newtonian potential

$$V = -G\sum_{k<l} \frac{m_k m_l}{r_{kl}}, \tag{3.98}$$

where $r_{kl} = |\mathbf{x}_k - \mathbf{x}_l|$ is the relative distance (more generally, one can take any potential $V(r_{kl})$). With this information at hand, one can construct the following action:

$$S[\mathbf{x}_k(t)] = 2\int \mathrm{d}\tau\ \sqrt{-VT}, \tag{3.99}$$

which is homogeneous in the velocities $\mathrm{d}\mathbf{x}/\mathrm{d}\tau$ and therefore reparametrization-invariant with respect to τ; cf. Section 3.1.1. After the horizontal stacking is performed, one is in a preferred frame in which the intrinsic velocities coincide with the ordinary velocities.

The equations of motion constructed from the action (3.99) read

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\sqrt{\frac{-V}{T}}m_k\frac{\mathrm{d}\mathbf{x}_k}{\mathrm{d}\tau}\right) = -\sqrt{\frac{T}{-V}}\frac{\partial V}{\partial \mathbf{x}_k}. \tag{3.100}$$

Note that this is a *non-local* equation because the frame is determined by the global stacking procedure, and, also, the total kinetic and potential energy of the universe occur explicitly. The gauge invariance with respect to translations and rotations leads to the constraints

$$\mathbf{P} = \sum_k \mathbf{p}_k = 0\ , \quad \mathbf{L} = \sum_k \mathbf{x}_k \times \mathbf{p}_k = 0, \tag{3.101}$$

that is, the total momentum and angular momentum of the universe are constrained to vanish. Since the momentum of the kth particle is given by

$$\mathbf{p}_k = \frac{\partial L}{\partial \dot{\mathbf{x}}_k} = m_k\sqrt{\frac{-V}{T}}\dot{\mathbf{x}}_k,$$

one finds the Hamiltonian constraint

$$H \equiv \sum_{k=1}^{n} \frac{\mathbf{p}_k^2}{2m_k} + V = 0, \tag{3.102}$$

which is a consequence of reparametrization invariance with respect to τ; see Section 3.1.1. Equation (3.100) can be drastically simplified if a convenient gauge choice is made for τ: it is chosen such as to make the total energy vanish,

$$T + V = 0. \tag{3.103}$$

Note that (3.103) is not the usual energy equation, since there is no external time present here. On the contrary, this equation is used to *define* time.

With (3.103), one then gets just Newton's equations from (3.100). Therefore, only after this choice has been made is the connection with Newtonian mechanics established. The in-principle observational difference from Newtonian mechanics is that here the total energy, momentum, and angular momentum must vanish. We note in this connection that in 1905 Henri Poincaré argued for a definition of time that makes the equations of motion assume their simplest form. He writes[5]

> Time must be defined in such a way that the equations of mechanics are as simple as possible. In other words, there is no way to measure time that is more true than any other; the one that is usually adopted is only more *convenient*.

It is, however, a fact that the choice (3.103) is distinguished not only because then the equations of motion (3.100) take their simplest form but also because only such a choice will ensure that the various clocks of (approximately isolated) subsystems

[5] 'Le temps doit être défini de telle façon que les équations de la mécanique soient aussi simples que possible. En d'autres termes, il n'y a pas une manière de mesurer le temps qui soit plus vraie qu' une autre; celle qui est généralement adoptée est seulement plus *commode*.' (Poincaré 1970)

march in step, since $\sum_k (T_k + V_k) = \sum_k E_k = 0$ (Barbour 1994, 2009). The only truly isolated system is the universe as a whole, and to determine time it is (in principle) necessary to monitor the whole universe. In practice, this is done even when atomic clocks are employed, for example in the determination of the arrival times of pulses from binary pulsars (Damour and Taylor 1991).

In this approach, the inertial frame and absolute time of Newtonian mechanics are *constructed* from observations through the minimization of the kinetic energy and the above choice of τ. One could call this a Leibnizian or Machian point of view. The operational time defined by (3.103) corresponds to the notion of 'ephemeris time' used in astronomy. That time must be defined such that the equations of motion should be simple was already known by Ptolemy (Barbour 1989). His theory of the Moon only took a simple form if sidereal time (defined by the rotation of the heavens, i.e. the rotation of the Earth) was used. This choice corresponds to a 'uniform flow of time'.

Time-reparametrization-invariant systems have already been discussed in Section 3.1.1 in connection with the parametrized non-relativistic particle. In contrast to there, however, *no* absolute time is present here and the theory relies exclusively on observational elements.

A formal analogy of the action (3.99) is given by the Jacobi action[6] in classical mechanics (Barbour 1986; Lanczos 1986; Brown and York 1989). Denoting the total (conserved) energy by E, the Jacobi action can be written in the form

$$S_J[\mathbf{x}_k(t)] = 2\int_A^B \mathrm{d}t\ \sqrt{(E-V)T}, \tag{3.104}$$

where the path leads from A to B, and t is Newton's absolute time. Using the expression (3.97) for the kinetic energy (with t instead of τ) and

$$\mathrm{d}s^2 := \sum_{k=1}^{n} m_k\, \mathrm{d}\mathbf{x}_k\, \mathrm{d}\mathbf{x}_k,$$

where s parametrizes the paths in configuration space, one gets the timeless form of the Jacobi action,

$$S_J[\mathbf{x}_k(t)] = \int \mathrm{d}s\ \sqrt{2(E-V)}. \tag{3.105}$$

Writing $\mathrm{d}s = \dot{s}\,\mathrm{d}t$, one has $\dot{s} = \sqrt{2T}$; if $E = 0$, then Jacobi's action is proportional to the action (3.99). We remark that (3.105) is also the expression that appears in the exponent of the WKB wave function Ψ for the stationary Schrödinger equation,

$$\psi \propto \frac{1}{|E-V|^{1/4}} \mathrm{e}^{\pm(\mathrm{i}/\hbar)\int \mathrm{d}s\,\sqrt{2m(E-V)}}.$$

The 'timeless' description (3.105) of mechanics employs only paths in configuration space. A 'speed' is determined later by solving the energy equation $T + V = E$ (T contains the velocities $\dot{\mathbf{x}}_k$). Barbour (1986) argues that for an isolated system, such

[6]There is a close analogy of this formulation with Fermat's principle of least time in geometrical optics.

as the universe is assumed to be, this demonstrates the redundancy of the notion of an independent time. All the essential dynamical content is already contained in the timeless paths. We shall see in Chapter 4 that GR can be described by an action similar to (3.99) and thus can be interpreted as 'already parametrized'. In this sense, the constraints of relational mechanics correspond to those of GR (see Chapter 4) more closely than the other examples considered in this chapter and suggest that GR is timeless in a significant sense.

Quantization of the Barbour–Bertotti model then follows Dirac's prescription, leading to the quantized constraints

$$\hat{H}\psi(\mathbf{x}_k) = 0 \ , \quad \hat{\mathbf{P}}\psi(\mathbf{x}_k) = 0 \ , \quad \hat{\mathbf{L}}\psi(\mathbf{x}_k) = 0. \tag{3.106}$$

Due to the latter two conditions, the wave function is actually defined on the relative configuration space, that is, the space of relative distances. We shall see in Section 4.2.3 how far the analogy between this model and GR reaches.

3.5 General remarks on constrained systems

Constrained systems play a central role in physics. The strong and electromagnetic interactions are described by gauge theories, which belong to this class. Another important example is GR. Since the quantization of GR is a non-trivial application of the quantization of constrained systems, the formalism developed for such systems is needed in the discussion of quantum gravity at several places. For this reason, we shall recapitulate here some of the basic properties of constrained systems. A detailed treatment can be found in many reviews and textbooks; see, for example, Dirac (1964), Hanson *et al.* (1976), Henneaux and Teitelboim (1992), Sundermeyer (1982), and Trautmann (1962). The pioneering paper on this subject was written by Léon Rosenfeld, who at that time was an assistent to Pauli in Zürich and was explicitly interested in the quantization of GR (Rosenfeld 1930*b*). Historical remarks can be found, for example, in DeWitt (1967*a*), Bergmann (1989), and Rovelli (2004). Most of the terminology that is used today stems from the work by Dirac and the group around Bergmann.

Here, we are mainly interested in the relation between constraints and invariances. We start with a brief discussion of Noether's theorems and then establish the connection to constrained systems.

Let us consider for simplicity a system of N particles in mechanics described by their positions $\{q^i(t)\}$, $i = 1, \ldots, N$. Their dynamics is given by the action

$$S[q^i(t)] = \int_{t_1}^{t_2} \mathrm{d}t \ L(q^i, \dot{q}^i). \tag{3.107}$$

We now assume that the dynamics is invariant under

$$q^i(t) \to q^i(t) + \delta q^i(t),$$

where

$$\delta q^i(t) = \epsilon^a(t) F^i_a(q, \dot{q}) + \dot{\epsilon}^a(t) G^i_a(q, \dot{q}), \tag{3.108}$$

with $a = 1, \ldots, n$.[7] The $\{\epsilon^a(t)\}$ are arbitrary functions of time, except possibly at the end points. If we set $q^0 \equiv t$, (3.108) can also include the case of time-reparametrization invariance discussed in Section 3.1.1. Under (3.108), the action (3.107) changes by

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \; \delta q^i L_i + \frac{\partial L}{\partial \dot{q}^i} \delta q^i \bigg|_{t_1}^{t_2}, \tag{3.109}$$

where

$$L_i := \frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}^i} \tag{3.110}$$

are the Euler derivatives. The equations of motion are satisfied if and only if $L_i = 0$. Using (3.108) in (3.109), one gets

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \; \epsilon^a \left[F_a^i L_i - \frac{\mathrm{d}}{\mathrm{d}t}(G_a^i L_i) \right] + \epsilon^a \left[F_a^i \frac{\partial L}{\partial \dot{q}^i} + G_a^i L_i \right]_{t_1}^{t_2} + \dot{\epsilon}^a G_a^i \frac{\partial L}{\partial \dot{q}^i} \bigg|_{t_1}^{t_2}.$$

Invariance of the action up to boundary terms then demands that

$$F_a^i L_i - \frac{\mathrm{d}}{\mathrm{d}t}(G_a^i L_i) = 0. \tag{3.111}$$

These n identities are known as generalized Bianchi identities or Noether's second theorem (Noether 1918). They are trivially satisfied for $L_i = 0$, that is, when the equations of motion are satisfied, but hold irrespective of them.

An important special case is $\epsilon^a(t) = \epsilon^a = \text{const.}$, that is, $\delta q^i = \epsilon^a F_a^i$. One then has

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \; \epsilon^a F_a^i L_i + \epsilon^a F_a^i p_i \bigg|_{t_1}^{t_2},$$

where the canonical momentum p_i has been introduced.The invariance of S up to boundary terms then leads to $L_i = 0$ instead of (3.111), that is, to the condition that the equations of motion must be satisfied. Defining $F_a^i p_i =: C_a$, the demand for invariance of the action *including* the boundary terms leads to

$$F_a^i p_i - C_a \bigg|_{t_2} = F_a^i p_i - C_a \bigg|_{t_1}.$$

An invariance under a group with a finite number of parameters thus leads to *conservation laws*. This is also called Noether's first theorem. We shall see that the above more general invariance under a group parametrized by functions of time leads instead to *constraints*.

In her pioneering paper (Noether 1918, p. 240), Noether writes[8]:

[7] In field theory, one would assume an invariance under $\delta\phi^A(x) = \epsilon^a(x) F_a^A + \epsilon^a,\mu(x) G_a^{A\mu}$; cf. Sundermeyer (1982).

[8] '...enthält Satz I alle in Mechanik u.s.w. bekannten Sätze über erste Integrale, während Satz II als größtmögliche gruppentheoretische Verallgemeinerung der "allgemeinen Relativitätstheorie" bezeichnet werden kann.'

...Theorem I contains all theorems about first integrals known in mechanics etc., while Theorem II can be considered as the largest possible group theoretical generalization of the 'general theory of relativity'.

Refering to remarks made by Hilbert, she writes at the end of her paper[9]:

Hilbert expresses his statement in the form that the failure of proper energy laws is a characteristic feature of the 'general theory of relativity'. In order for this statement to be literally valid, one must thus interpret the name 'general relativity' more generally than usual and also extend it to the above groups that depend on n arbitrary functions.

In a final footnote, she emphasizes the correctness of a remark made by Felix Klein eight years earlier, who said that the usual terminology of 'relativity' in physics should be replaced by 'invariance with respect to a group'.

Let us return to the identities (3.111). The Euler derivatives (3.110) can be written as (see e.g. Sundermeyer 1982)

$$L_i = \frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\dot{q}^j - \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\ddot{q}^j =: V_i - W_{ij}\ddot{q}^j = 0.$$

Inserting this into (3.111), one finds

$$L_i(F^i_a - \dot{G}^i_a) - G^i_a \dot{V}_i + G^i_a \dot{W}_{ij}\ddot{q}^j + G^i_a W_{ij}\dddot{q}^j.$$

Since $\dddot{q}^j$ only occurs in the last term, its coefficient must vanish separately:

$$G^i_a W_{ij} = 0.$$

If at least one of the G_a is non-vanishing, the non-vanishing set consists of null eigenvectors of the Hessian W_{ij}. If all the G_a are vanishing (and if the ϵ^a are not constant), one has $L_i F^i_a = 0 = V_i - W_{ij}\ddot{q}^j$ and thus, from the separate vanishing of the second-derivative term, the condition $F^i_a W_{ij} = 0$; again, the Hessian has null eigenvectors. A Lagrangian with this property is called *singular*. The invariance under (3.108) thus has an important consequence: one cannot isolate the accelerations in the equation of motion, and the usual determination of the trajectories by specifying initial conditions breaks down. The degeneracy of the Hessian also means that one cannot apply the canonical formalism in the usual way, since not all velocities can be replaced by momenta. As a consequence, there remain some constraints between the q^i and the p_i. We thus see that the presence of invariances leads to constraints (the opposite does not hold in general).

The application of the canonical formalism to a constrained system is the content of the Dirac–Bergmann algorithm (which would, perhaps, more properly be called the Rosenfeld–Dirac–Bergmann algorithm), which is treated at length in the references cited at the beginning of this section. We summarize here the essential features that are relevant to quantum gravity.

[9]'Hilbert spricht seine Behauptung so aus, daß das Versagen eigentlicher Energiesätze ein charakteristisches Merkmal der "allgemeinen Relativitätstheorie" sei. Damit diese Behauptung wörtlich gilt, ist also die Bezeichnung "allgemeine Relativität"weiter als gewöhnlich zu fassen, auch auf die vorangehenden von n willkürlichen Funktionen abhängenden Gruppen auszudehnen.'

The degeneracy of the Hessian leads to a set of constraints,

$$\phi_\mu(q,p) = 0, \tag{3.112}$$

where $\mu = 1, \ldots, n$. They are called *primary constraints* because they follow directly from the form of the Lagrangian before use is made of the equations of motion. We shall occasionally also use Dirac's notion of 'weak equality' (Dirac 1964) and write $\phi_\mu(q,p) \approx 0$ instead of (3.112). This notation is used to emphasize that the constraints should be imposed only after all Poisson brackets between phase-space functions are evaluated.

Calling the usual canonical Hamiltonian H_c, the equations of motion are found to contain arbitrary functions $v^\mu(t)$:

$$\begin{aligned}\dot{q}^i &= \{q^i, H_c\} + v^\mu\{q^i, \phi_\mu\} \equiv \{q^i, H_T\},\\ \dot{p}_i &= \{p_i, H_c\} + v^\mu\{p_i, \phi_\mu\} \equiv \{p_i, H_T\}.\end{aligned} \tag{3.113}$$

In these equations, we have introduced

$$H_T := H_c + v^\mu \phi_\mu, \tag{3.114}$$

which was called the 'total Hamiltonian' by Dirac (and later frequently called the 'Dirac Hamiltonian'). The consistency of the formalism requires that the primary constraints $\phi_\mu = 0$ be conserved in time; that is, one must have $\{\phi_\mu, H_T\} \approx 0$. In this way, some of the functions $v^\mu(t)$ may be fixed, while others remain arbitrary. It can also happen (and this is usually the case) that further constraints arise from this demand. These are called *secondary constraints.*

An important classification of the constraints is, in Dirac's terminology, the distinction between first- and second-class constraints. First-class constraints are defined by the Poisson-bracket relations

$$\{\phi_a, \phi_b\} = f^c_{ab}\phi_c\ ; \tag{3.115}$$

that is, they close among themselves. Constraints which do not obey these relations are called *second-class constraints.* They play a role, for example, in supergravity; cf. Section 5.3.6.

The consistency conditions lead in a finite number of steps to a set of constraints, some of which are of first and some of second class. The functions v^μ associated with second-class constraints become fixed, while the functions associated with first-class constraints remain arbitrary. The result of the algorithm is then the following equations of motion:

$$\begin{aligned}\dot{q}^i &= \{q^i, H'\} + v^a\{q^i, \phi_a\},\\ \dot{p}_i &= \{p_i, H'\} + v^a\{p_i, \phi_a\},\end{aligned} \tag{3.116}$$

where ϕ_a denotes the final set of *primary first-class* constraints. Of course, the whole set of constraints, denoted by $\phi_A = 0$ (of which the ϕ_a occuring in (3.116) are only a subset), must be fulfilled. In (3.116), H' denotes the final 'first-class Hamiltonian',[10] which differs from H_c by $v^i\phi_i$, where the $\phi_i = 0$ are the second-class constraints and the v^i are fixed functions; in GR, we shall have $H' = H_\mathrm{c}$ and this distinction is not needed. We can formulate the set of equations (3.116) similarly to (3.113) by again introducing a total Hamiltonian as in (3.114); this time, however, $H_\mathrm{T} := H' + v^a\phi_a$.

The existence of the arbitrary functions v^a in (3.116) signals the presence of redundancies. The corresponding transformations are often called 'gauge transformations', although this terminology was not used by Dirac in this context. In fact, gauge theories are only particular examples of constrained systems.[11] It is a most important result of the Dirac–Bergmann algorithm that first-class constraints are generators of redundancy ('gauge') transformations; see, for example, the careful discussion in Pons (2005). These redundancy transformations map entire solutions of the equations of motion into entire solutions; points on one trajectory are mapped to points on another trajectory at the same time t. If v and v' are used as shorthands for the arbitrary functions associated with any two trajectories, then this map from one trajectory $g_v(t)$ to any other trajectory $g_{v'}(t)$ is characterized by

$$\Delta g(t) := g_{v'}(t) - g_v(t) = \{g, G(t)\}_{g_v}, \tag{3.117}$$

where $G(t) \equiv G(q, p, t)$ is a phase-space function that must preserve the constraints and thus must be of first class. In fact, $G(t)$ is of the following form (Castellani 1982):

$$G(t) = G_0\xi(t) + G_1\dot{\xi}(t) + G_2\ddot{\xi}(t) + \ldots = \sum_{i=0}^{N} G_i\xi^{(i)}(t), \tag{3.118}$$

where $\xi(t)$ is an arbitrary function of t, and the G_i are phase-space functions to be determined for the system under investigation. It turns out that G_N is a primary first-class constraint, while $G_{N-1}, \ldots, G_0$ are secondary first-class constraints. That is, the generator of redundancy transformations is a particular linear combination of the first-class constraints. As an example, we mention vacuum electrodynamics, where there are two constraints: the primary first-class constraint $G_1 = p^0$, and the second-class constraint $G_0 = \nabla\mathbf{E}$ (Gauss's law), which together generate the U(1) gauge transformations of electrodynamics. For us, the most important example is GR, which is characterized by the presence of Hamiltonian and momentum constraints; this is the subject of Chapter 4.

In the total Hamiltonian H_T, only the primary first-class constraints occur. Since all first-class constraints generate redundancy transformations, Dirac conjectured that *all* first-class constraints should be considered in the definition of the total Hamiltonian (Dirac 1964). This, however, will not be done here, since it is neither demanded by nor in conflict with the formalism (Pons 2005).

[10] A first-class function is characterized by the fact that its Poisson brackets with the constraints are linear combinations of the constraints.

[11] A gauge theory is usually defined as follows. Start from a rigid symmetry with a conserved current. Making this symmetry local requires the introduction of a new field, the 'gauge field'. Adding a kinetic term for this new field makes it dynamical. The paradigmatic example is the Yang–Mills field A^i_μ; cf. Section 4.1.3.

Functions $A(q,p)$ for which $\{A, \phi_A\} \approx 0$ holds are often called *observables* because they do not change under a redundancy transformation. It must be emphasized that there is no a priori relation of these observables to observables in an operational sense. This notion was introduced by Bergmann in the hope that these quantities might play the role of the standard observables in quantum theory (Bergmann 1961).

In order to select one physical representative from amongst all equivalent configurations, one frequently employs 'gauge conditions' in order to break the redundancies. Although gauge theories are, as mentioned above, only special examples of constrained systems, it is customary to talk here about gauge freedom and gauge fixing. An important application is in path-integral quantization; see Section 2.2.3. A 'gauge' should be chosen in such a way that there is no further gauge freedom left and that any configuration can be transformed into one satisfying the gauge. The first condition is sometimes violated ('Gribov ambiguities'), but this is irrelevant for infinitesimal gauge transformations. If one identifies all points on the same gauge orbit, one arrives at the *reduced phase space* of the theory. In the general case, the reduced phase space is not a cotangent bundle; that is, one cannot identify which variables are the q's and which are the p's.

Instead of gauge fixing, one can keep the redundancies in the classical theory and perform a quantization by implementing the constraints in the way done in (3.13).

Further reading: Barbour (2009), Kuchař (1973), Sundermeyer (1982).

4

Hamiltonian formulation of general relativity

4.1 The seventh route to geometrodynamics

It is the purpose of this chapter to develop the Hamiltonian formulation of general relativity, which will serve as the starting point for quantization in Chapters 5 and 6. In the present section, we shall derive it directly from the algebra of surface deformations (Section 3.3), while in the next section it will be recovered from the Einstein–Hilbert action through a 3+1 decomposition. In these two sections, the formalism will be applied to the traditional metric formulation, while in the last section of the chapter a more recent formalism using connections will be used. The quantization of the latter version will be treated in Chapter 6.

That the theory of GR in its Hamiltonian version can be constructed directly from the algebra of surface deformations (3.84)–(3.86) was shown by Hojman *et al.* (1976), and we shall follow their exposition in this section. Hojman *et al.* call their approach the 'seventh route to geometrodynamics', complementing the six routes presented in Box 17.6 of Misner *et al.* (1973).[1] In Hojman *et al.* (1976), a four-dimensional space–time is presupposed from the outset. There is an alternative approach in which only spatial three-geometries are presupposed and embeddability is not needed, see Barbour *et al.* (2002) and Anderson (2007*a*).

4.1.1 Principle of path independence

The starting point is the assumption that the *only* gravitational canonical pair of variables on the spatial hypersurfaces consists of the three-dimensional metric $h_{ab}(\mathbf{x})$ and its conjugate momentum $p^{cd}(\mathbf{y})$. In addition, matter fields may be present. The constraints will not be imposed—in fact, they will be *derived* from the algebra. Different classical theories of gravity typically contain additional degrees of freedom in the configuration space of the gravitational sector (such as a scalar field in the Brans–Dicke case or the extrinsic curvature for R^2-theories). The central idea in the derivation is the use of a 'principle of path independence'. Let us assume the presence of two different three-dimensional geometries (three-geometries for short) $\mathcal{G}_{\text{in}}$ and $\mathcal{G}_{\text{fin}}$ and a set of observers distributed on $\mathcal{G}_{\text{in}}$; the observers bifurcate and follow different evolutions of intermediate hypersurfaces, making records of them as well as of the fields on them, until they all end up on $\mathcal{G}_{\text{fin}}$. The principle of path independence states that the change

[1]For more routes, see Section 3 in Anderson (2007*a*).

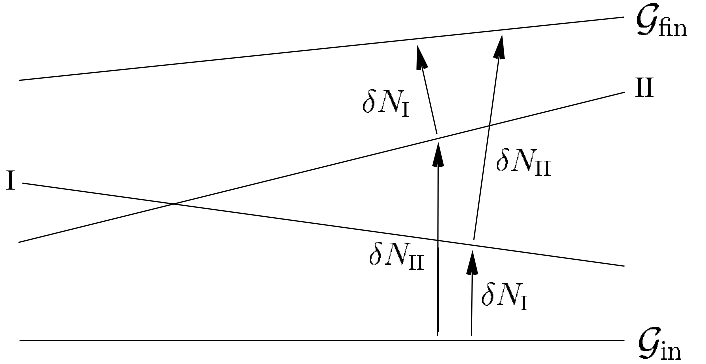

Fig. 4.1 Normal deformations along two different intermediate three-geometries.

in all field variables must be independent of the route that is chosen between $\mathcal{G}_{\text{in}}$ and $\mathcal{G}_{\text{fin}}$. Only in this case can the evolution of three-geometries be interpreted as arising from different foliations through the same four-dimensional space–time.[2]

The evolution of a function F of the gravitational canonical variables is given by (cf. (3.83))

$$\begin{aligned}\dot{F}(h_{ab}(x), p^{cd}(x)) &= \int \mathrm{d}x' \ (\{F, \mathcal{H}_\perp(x')\}N(x') + \{F, \mathcal{H}_a(x')\}N^a(x')) \\ &\equiv \int \mathrm{d}x' \, \{F, \mathcal{H}_\mu(x')\}N^\mu(x'),\end{aligned} \tag{4.1}$$

where $\mathrm{d}x'$ is a shorthand for d^3x', and x for $\mathbf{x}$, etc. Consider now an infinitesimal evolution along two different intermediate hypersurfaces (Fig. 4.1): first, a normal deformation from $\mathcal{G}_{\text{in}}$ to I and then from I to $\mathcal{G}_{\text{fin}}$; and second, a normal deformation from $\mathcal{G}_{\text{in}}$ to II and then from II to $\mathcal{G}_{\text{fin}}$. Evolving F along the first possibility yields

$$\begin{aligned}F\left[\mathcal{G}_{\text{fin}}\right] = F\left[\mathcal{G}_{\text{in}}\right] &+ \int \mathrm{d}x' \ \{F, \mathcal{H}_\perp(x')\}\delta N_{\text{I}}(x') \\ &+ \int \mathrm{d}x' \ \{F, \mathcal{H}_\perp(x')\}\delta N_{\text{II}}(x') \\ &+ \int \mathrm{d}x' \, \mathrm{d}x'' \ \{\{F, \mathcal{H}_\perp(x')\}\delta N_{\text{I}}(x'), \mathcal{H}_\perp(x'')\} \, \delta N_{\text{II}}(x'').\end{aligned}$$

The first two terms on the right-hand side describe the evolution of F from $\mathcal{G}_{\text{in}}$ to I, and the last two terms the evolution from I to $\mathcal{G}_{\text{fin}}$. The option of reaching $\mathcal{G}_{\text{fin}}$ via hypersurface II gives an analogous result. Taking the difference between the two expressions and using the Jacobi identity leads one to

$$\delta F = -\int \mathrm{d}x' \, \mathrm{d}x'' \ \{\{\mathcal{H}_\perp(x'), \mathcal{H}_\perp(x'')\}, F\} \, \delta N_{\text{I}}(x'')\delta N_{\text{II}}(x'). \tag{4.2}$$

[2]This holds only for cases where the sandwich conjecture is satisfied, that is, for cases in which two three-geometries uniquely determine the space–time 'sandwiched' between them.

From (3.91) or (3.88), one knows that the difference between these two normal deformations must be given by a tangential deformation,[3]

$$\delta F = -\int \mathrm{d}x' \; \{F, \mathcal{H}_a(x')\}\delta N^a(x'), \tag{4.3}$$

where

$$\delta N^a = -\sigma h^{ab} \left(\delta N_{\mathrm{I}} \delta N_{\mathrm{II},b} - \delta N_{\mathrm{II}} \delta N_{\mathrm{I},b}\right). \tag{4.4}$$

Inserting for calculational purposes a delta function into (4.3) and performing a partial integration, one has

$$\begin{aligned}
&-\int \mathrm{d}x' \; \{F, \mathcal{H}_a(x')\}\delta N^a(x') = \\
&-\sigma \int \mathrm{d}x' \, \mathrm{d}x'' \; \frac{\partial}{\partial x'^b}\delta(x'-x'')\{F, \mathcal{H}_a(x'')\}h^{ab}(x'')\delta N_{\mathrm{I}}(x'')\delta N_{\mathrm{II}}(x') \\
&-\sigma \int \mathrm{d}x' \, \mathrm{d}x'' \; \frac{\partial}{\partial x'^b}\delta(x'-x'')\{F, \mathcal{H}_a(x')\}h^{ab}(x')\delta N_{\mathrm{I}}(x'')\delta N_{\mathrm{II}}(x').
\end{aligned} \tag{4.5}$$

Setting this equal to (4.2) and using the arbitrariness of $\delta N_{\mathrm{I}}(x'')\delta N_{\mathrm{II}}(x')$, one finds

$$\begin{aligned}
\{F, \{\mathcal{H}_\perp(x'), \mathcal{H}_\perp(x'')\}\} = &-\sigma \delta_{,b}(x'-x'')h^{ab}(x'')\{F, \mathcal{H}_a(x'')\} \\
&-\sigma \delta_{,b}(x'-x'')h^{ab}(x')\{F, \mathcal{H}_a(x')\}.
\end{aligned} \tag{4.6}$$

Inserting the Poisson bracket (3.84) on the left-hand side, one finds the condition

$$\frac{\partial}{\partial x'^a}\delta(x'-x'')\left(\{F, h^{ab}(x')\}\mathcal{H}_b(x') + \{F, h^{ab}(x'')\}\mathcal{H}_b(x'')\right) = 0. \tag{4.7}$$

As this should hold for all F, the generators $\mathcal{H}_a$ themselves must vanish as constraints, $\mathcal{H}_a \approx 0$. This result from the principle of path independence only follows because the Poisson bracket (3.84) depends on the metric h^{ab}. Since $\mathcal{H}_a \approx 0$ must hold on every hypersurface, it must be conserved under a normal deformation. From (3.85) one then finds that $\mathcal{H}_\perp$ must also vanish, $\mathcal{H}_\perp \approx 0$. We have thus shown that the algebra of surface deformations, together with the principle of path independence (equivalent to the principle of embeddability), enforces the constraints

$$\mathcal{H}_\perp \approx 0\,, \quad \mathcal{H}_a \approx 0. \tag{4.8}$$

This result follows for any number of space–time dimensions except two. For two dimensions (after a suitable definition of the $\mathcal{H}_\mu$), the metric does not appear on the right-hand side of the algebra (Teitelboim 1984). This leads to the possibility of having path independence without constraints, resulting in potential Schwinger terms in the quantum theory (cf. Sections 3.2 and 5.3.5).

[3]Recall that $\mathcal{H}_a$ generates *minus* the surface deformations.

4.1.2 Explicit form of generators

How does one find the explicit form of $\mathcal{H}_\perp$ and $\mathcal{H}_a$? The constraints $\mathcal{H}_a$ generate coordinate transformations on a hypersurface. If the transformation law of certain fields is given, $\mathcal{H}_a$ follows (or, conversely, the transformation properties are determined by a given $\mathcal{H}_a$). Knowing $\mathcal{H}_a$, one can consider the Poisson bracket (3.84) as a system of infinitely many equations to determine $\mathcal{H}_\perp$, which must then depend on h_{ab} because the right-hand side of (3.84) does. If, moreover, the assumption is made that $\mathcal{H}_\perp$ depends *only* on h_{ab} (and its momentum) as gravitational degrees of freedom, GR will follow.

Some general properties of $\mathcal{H}_a$ can be inferred without studying the change of different fields under coordinate transformations, see Teitelboim (1980) for proofs and more details. The general form of $\mathcal{H}_a$ can be restricted by the following two requirements:

1. It must be linear in the momenta in order to generate transformations of the fields under coordinate transformations (and not mix fields and momenta).
2. It should contain the momenta only up to the first spatial derivatives because it should generate first-order derivatives in the fields (Taylor expansion to first order).

Therefore,

$$\mathcal{H}_a = b_a{}^b{}_B(\phi_C)p^B{}_{,b} + a_{aB}(\phi_C)p^B, \tag{4.9}$$

where ϕ_A is now a symbolic notation for all fields, including gravity, and p^A denotes the corresponding momenta. If one demands the existence of ultralocal solutions to (3.84) (i.e. a deformation localized at a point $\mathbf{x}_0$ can change the field only at $\mathbf{x}_0$), one finds that

$$b^{ab}_B = b^{ba}_B. \tag{4.10}$$

General requirements allow the form of $\mathcal{H}_\perp$ to be restricted as well (Teitelboim 1980). Writing $\mathcal{H}_\perp$ as the sum of a gravitational and a matter part (with 'matter' referring here to all non-gravitational bosonic fields,[4] symbolically denoted by ϕ),

$$\mathcal{H}_\perp = \mathcal{H}_\perp^{\mathrm{g}}[h_{ab}, p^{cd}] + \mathcal{H}_\perp^{\mathrm{m}}[h_{ab}, p^{cd}; \phi, p_\phi], \tag{4.11}$$

it follows from (3.84) that $\mathcal{H}_\perp^{\mathrm{m}}$ *must* depend on the gravitational degrees of freedom; that is, gravity couples to all forms of matter. This is the Hamiltonian version of Weinberg's result discussed in Section 2.1. More restrictions can be obtained if one assumes that $\mathcal{H}_\perp^{\mathrm{m}}$ does not depend on the gravitational momenta p^{cd}. This corresponds to the presence of 'non-derivative couplings' only; that is, there are no gravitational velocities on the Lagrangian level and therefore there is no modification of the relationship between momentum and velocity. Then one can show that $\mathcal{H}_\perp^{\mathrm{m}}[h_{ab}; \phi, p_\phi]$ depends on the h_{ab} only ultralocally; that is, no derivatives or integrals of h_{ab} appear. From this one can infer that the terms $\mathcal{H}_\perp^{\mathrm{m}}$ obey the relation (3.84) separately, that is,

$$\{\mathcal{H}_\perp^{\mathrm{m}}(x), \mathcal{H}_\perp^{\mathrm{m}}(y)\} = -\sigma\delta_{,a}(x,y)\left(h^{ab}(x)\mathcal{H}_b^{\mathrm{m}}(x) + h^{ab}(y)\mathcal{H}_b^{\mathrm{m}}(y)\right). \tag{4.12}$$

One also finds from the demand for ultralocality that $\mathcal{H}_\perp$ must depend ultralocally on the momenta.

[4] For the fermionic case, see the remarks below.

Using these general properties, the behaviour of various fields under coordinate transformations generated by $\mathcal{H}_a$ can be studied and the detailed form of $\mathcal{H}_a$ and $\mathcal{H}_\perp$ be derived. The first case is a scalar field ϕ. Under an infinitesimal coordinate transformation $x'^a = x^a - \delta N^a(x)$, it transforms as

$$\delta\phi(x) := \phi'(x) - \phi(x) \approx \frac{\partial \phi}{\partial x^a}\delta N^a \equiv \mathcal{L}_{\delta\mathbf{N}}\phi, \tag{4.13}$$

where $\mathcal{L}$ denotes the Lie derivative. This is generated by

$$\mathcal{H}_a = p_\phi \phi_{,a}. \tag{4.14}$$

Comparison with (4.9) shows that $b^{ab} = 0 = b^{ba}$, so that (4.10) is fulfilled and ultralocality holds.

For a vector field $A_a(x)$, the transformation is

$$\delta A_a = A_{a,b}\delta N^b + A_b \delta N^b{}_{,a} \equiv (\mathcal{L}_{\delta\mathbf{N}}A)_a\,, \tag{4.15}$$

which is generated by

$$\mathcal{H}_a = -p^b{}_{,b}A_a + (A_{b,a} - A_{a,b})p^b. \tag{4.16}$$

Comparison with (4.9) shows that

$$b_a{}^b{}_C = -A_a \delta^b{}_C. \tag{4.17}$$

Therefore, the condition for ultralocality (4.10) is not fulfilled for vector fields. Its restoration will lead to the concept of gauge theories (Section 4.1.3).

For a covariant tensor field of second rank (not necessarily symmetric) $t_{ab}(x)$, one has

$$\delta t_{ab} = t_{ab,c}\delta N^c + t_{ac}\delta N^c{}_{,b} + t_{cb}\delta N^c{}_{,a} \equiv (\mathcal{L}_{\delta\mathbf{N}}t)_{ab}\,, \tag{4.18}$$

which is generated by

$$\mathcal{H}_a = t_{bc,a}p^{bc} - (t_{ab}p^{cb})_{,c} - (t_{ca}p^{cb})_{,b}. \tag{4.19}$$

It turns out that in order for (4.10) to be fulfilled, one must have

$$t_{ab} = f(x)h_{ab} \tag{4.20}$$

with an arbitrary function $f(x)$; that is, the tensor field must be proportional to the metric itself. Choosing in particular $t_{ab} = h_{ab}$, one finds for the generator (4.19) the expression

$$\mathcal{H}_a^{\mathrm{g}} = -2p_a{}^c{}_{,c} + 2\Gamma^d_{ac}p_d{}^c \equiv -2D_b p_a{}^b \equiv -2p_a{}^b{}_{|b}. \tag{4.21}$$

The last two terms denote the covariant derivative in three dimensions (recall that p^{ab} is a tensor density of weight one).

Using the result (4.21) for $\mathcal{H}_a^{\mathrm{g}}$, one can construct from (3.84) an explicit expression for $\mathcal{H}_\perp^{\mathrm{g}}$. A rather lengthy but straightforward calculation leads to (Hojman *et al.* 1976; Teitelboim 1980)

$$\mathcal{H}_\perp^{\mathrm{g}} = 16\pi G\, G_{abcd} p^{ab} p^{cd} + V[h_{ab}], \tag{4.22}$$

with

$$G_{abcd} = \frac{1}{2\sqrt{h}} (h_{ac}h_{bd} + h_{ad}h_{bc} - h_{ab}h_{cd}) \tag{4.23}$$

as the (inverse) 'DeWitt metric',[5] h denoting the determinant of h_{ab}, and

$$V = \frac{\sigma\sqrt{h}}{16\pi G} ({}^{(3)}R - 2\Lambda), \tag{4.24}$$

where ${}^{(3)}R$ is the three-dimensional Ricci scalar.[6] The inverse of (4.23) is called the 'DeWitt' metric because it plays the role of a metric in the space of all metrics (DeWitt 1967*a*); cf. Section 4.2.5. For this reason, it is often referred to as the 'supermetric'. The explict expression reads

$$G^{abcd} = \frac{\sqrt{h}}{2} (h^{ac}h^{bd} + h^{ad}h^{bc} - 2h^{ab}h^{cd}) \tag{4.25}$$

(the last term here is the same in all space dimensions), obeying

$$G^{abcd} G_{cdef} = \tfrac{1}{2} \left(\delta^a_e \delta^b_f + \delta^a_f \delta^b_e \right). \tag{4.26}$$

We recall that the Poisson-bracket relation (3.85) states that $\mathcal{H}_\perp$ transforms as a scalar density under coordinate transformations; this is explicitly fulfilled by (4.22) (G_{abcd} has weight -1, and p^{ab} and V have weight 1, so $\mathcal{H}_\perp^{\mathrm{g}}$ has weight 1). We thus have

$$\delta\mathcal{H}_\perp^{\mathrm{g}}(x) = \int \mathrm{d}y\ \{\mathcal{H}_\perp^{\mathrm{g}}(x), \mathcal{H}_a^{\mathrm{g}}(y)\} \delta N^a(y) = \frac{\partial}{\partial x^a} \left(\mathcal{H}_\perp^{\mathrm{g}}(x) \delta N^a(x) \right). \tag{4.27}$$

It will be shown in Section 4.2 that $\mathcal{H}_\perp^{\mathrm{g}}$ and $\mathcal{H}_a^{\mathrm{g}}$ uniquely characterize GR, that is, they follow from the Einstein–Hilbert action (1.1). Finally, we want to remark that the uniqueness of the construction presented here ceases to hold in space dimensions greater than three (Teitelboim and Zanelli 1987).

4.1.3 Geometrodynamics and gauge theories

We have seen that for vector fields, $\mathcal{H}_a$ is of such a form that the condition of ultralocality for $\mathcal{H}_\perp$ would be violated; see (4.16). Since vector fields are an important ingredient in the description of nature, the question arises whether a different formulation can be found that is in accordance with ultralocality. For this purpose, it is

[5] In d space dimensions, the last term reads $-2/(d-1)h_{ab}h_{cd}$. We remark that there is a formal similarity between the DeWitt metric and the graviton propagator in the covariant approach, cf. (2.93).

[6] G and Λ are at this stage just free parameters. They will later be identified with the gravitational constant and the cosmological constant, respectively.

suggestive to omit the term $-p^b{}_{,b}A_a$ in (4.16) because then the $b_a{}^b{}_B$ will become zero. This leads to the replacement (Teitelboim 1980)

$$\mathcal{H}_a \to \bar{\mathcal{H}}_a := \mathcal{H}_a + p^b{}_{,b}A_a = (A_{b,a} - A_{a,b})p^b. \tag{4.28}$$

What happens with the Poisson-bracket relation (3.86) after this modification? A brief calculation shows

$$\{\bar{\mathcal{H}}_a(x), \bar{\mathcal{H}}_b(y)\} = \bar{\mathcal{H}}_b(x)\delta_{,a}(x,y) + \bar{\mathcal{H}}_a(y)\delta_{,b}(x,y) - F_{ab}(x)p^c{}_{,c}(x)\delta(x,y), \tag{4.29}$$

where $F_{ab} = \partial_a A_b - \partial_b A_a$. The new term in (4.29) will only be harmless if it generates physically irrelevant transformations ('gauge transformations'). This is the case if the new term actually vanishes as a constraint. Since setting F_{ab} to zero would appear too strong (leaving only the restricted option $A_a = \partial_a\varphi$), it is suggestive to demand that $p^a{}_{,a} \approx 0$. One therefore introduces the constraint

$$\mathcal{G}(x) := -\frac{1}{e}p^a{}_{,a}(x) \equiv -\frac{1}{e}E^a{}_{,a}(x) \equiv -\frac{1}{e}\nabla\mathbf{E}(x). \tag{4.30}$$

The constraint $\mathcal{G} \approx 0$ is just *Gauss's law* of electrodynamics (in the sourceless case), with the momentum being equal to the electric field $\mathbf{E}$ (the electric charge, e, has been introduced for convenience). As usual, Gauss's law generates gauge transformations,[7]

$$\delta A_a(x) = \int \mathrm{d}y\ \{A_a(x), \mathcal{G}(y)\}\xi(y) = \frac{1}{e}\partial_a\xi(x)\ , \tag{4.31}$$

$$\delta p^a(x) = \int \mathrm{d}y\ \{p^a(x), \mathcal{G}(y)\}\xi(y) = 0. \tag{4.32}$$

The electric field is of course gauge invariant, and so is the field strength F_{ab}. In the modified constraint (4.28), the first term ($\mathcal{H}_a$) generates the usual transformations for a vector field (see (4.15)), while the second term ($p^b{}_{,b}A_a$) generates gauge transformations for the 'vector potential' $A_a(x)$. Therefore, $A_a(x)$ transforms under $\bar{\mathcal{H}}_a$ not like a covariant vector but only like a vector modulo a gauge transformation. This fact has already been encountered in the space–time picture; see Section 2.1. The electric field, however, transforms as a contravariant tensor density, since the additional term in (4.28) has no effect.

The above introduction of the gauge principle can be extended in a straightforward manner to the non-Abelian case. Consider instead of a single $A_a(x)$ a set of several fields $A_a^i(x)$, $i = 1, \ldots, N$. The simplest generalization of the Abelian case consists in the assumptions that $A_a^i(x)$ should not mix with its momentum $p_i^a(x)$ under a gauge transformation, that the momenta should transform homogeneously, and that the gauge constraint (the non-Abelian version of Gauss's law) is local. This then leads to (Teitelboim 1980)

$$\mathcal{G}_i = -\frac{1}{f}p_i{}^a{}_{,a} + C_{ij}{}^k A_a{}^j p_k{}^a \approx 0, \tag{4.33}$$

[7]In order to generate the full U(1) transformations of electrodynamics, one also needs the constraint $p^0 \approx 0$, see Section 3.5.

where f and ${C_{ij}}^k$ are constants. Demanding that the commutation of two gauge transformations be again a transformation, it follows that the ${C_{ij}}^k$ are the structure constants of a Lie algebra. One then has

$$\{\mathcal{G}_i(x), \mathcal{G}_j(y)\} = {C_{ij}}^k \mathcal{G}_k(x)\delta(x,y), \tag{4.34}$$

which characterizes a *Yang–Mills theory*.

As in the case of the gravitational field, one can construct the corresponding part of the Hamiltonian constraint, $\mathcal{H}_\perp^{\mathrm{YM}}$, from the Poisson-bracket relation (3.84). Writing

$$\mathcal{H}_\perp = \mathcal{H}_\perp^{\mathrm{g}} + \mathcal{H}_\perp^{\mathrm{YM}} \tag{4.35}$$

and demanding that the Yang–Mills part be independent of the gravitational momenta (so that $\mathcal{H}_\perp^{\mathrm{YM}}$ depends only ultralocally on the metric and must therefore obey (3.84) separately), one is led to the form

$$\mathcal{H}_\perp^{\mathrm{YM}} = \frac{1}{2\sqrt{h}} \left(h_{ab}\gamma^{ij} p_i^a p_j^b - \sigma h^{ab} \gamma_{ij} B_a^i B_b^j \right), \tag{4.36}$$

where $\gamma_{ij} = {C_{ik}}^l {C_{jl}}^k$ is the 'group metric' (γ^{ij} being its inverse), and the $B_a^i = \frac{1}{2}\epsilon_{abc}F^{ibc}$ are the non-Abelian 'magnetic fields'. The non-Abelian field strength is given by

$$F_{ab}^i = \partial_a A_b^i - \partial_b A_a^i + f {C^i}_{jk} A_a^j A_b^k .$$

The Hamiltonian (4.36) can be found from the action

$$S_{\mathrm{YM}} = -\frac{\sigma}{4} \int \mathrm{d}^4 x \, \sqrt{-g} \gamma_{ij} F_{\mu\nu}^i F^{j\mu\nu}, \tag{4.37}$$

which is the usual Yang–Mills action.

To summarize, the principle of path independence together with the demand that $\mathcal{H}_\perp$ be ultralocal[8] in the momenta leads to the concept of gauge theories in a natural way.

What about fermionic fields? Recalling that the Dirac equation is the 'square root' of the Klein–Gordon equation, one may try to construct a similar 'square root' for the generators of surface deformations. This has been done by Tabensky and Teitelboim (1977); it leads to spin-3/2 fields and the concept of supergravity (cf. Section 2.3) but *not* to spin-1/2 fields. This could be a hint that the usual spin-1/2 fields only emerge through the use of superstrings (Chapter 9). The Hamiltonian formalism for supergravity will be discussed in Section 5.3.6.

4.2 The 3+1 decomposition of general relativity

It will be shown in this section that GR is characterized by having (4.21) and (4.22) as the constraints. This is achieved by choosing appropriate canonical variables and casting the Einstein–Hilbert action (1.1) into Hamiltonian form.

[8] As is known from the discussion of the Aharonov–Bohm effect, a formulation without the vector potential can only be obtained in a non-local way.

4.2.1 The canonical variables

The Hamiltonian formalism starts from the choice of a configuration variable and the definition of its momentum. Since the latter requires a time coordinate ('$p = \partial L/\partial \dot{q}$'), one must cast GR in a form where it exhibits a 'distinguished' time. This is achieved by *foliating* the space–time described by $(\mathcal{M}, g)$ into a set of three-dimensional space-like hypersurfaces Σ_t; cf. also Section 3.3. The covariance of GR is preserved by allowing for the possibility of considering *all* feasible foliations of this type.

This is of relevance not only for quantization (which is our motivation here), but also for important applications in the classical theory. For example, numerical relativity needs a description in terms of foliations in order to describe the dynamical evolution of events, for example the coalescence of black holes and their emission of gravitational waves (Baumgarte and Shapiro 2003).

As a necessary condition, we want to demand that $(\mathcal{M}, g)$ be globally hyperbolic, that is, it possesses a three-dimensional space-like Cauchy surface Σ (an 'instant of time') on which initial data can be prescribed to determine uniquely the whole space–time; see, for example, Wald (1984) or Hawking and Ellis (1973) for details. In such cases, the classical initial-value formulation makes sense, and the Hamiltonian form of GR can be constructed. The occurrence of naked singularities is prohibited by this assumption.

An important theorem states that for a globally hyperbolic space–time $(\mathcal{M}, g)$ there exists a global 'time function' f such that each surface $f =$ constant is a Cauchy surface; therefore, $\mathcal{M}$ can be foliated into Cauchy hypersurfaces, and its topology is a direct product between the real numbers $\mathbb{R}$ (representing 'time') and the three-dimensional manifold Σ (representing 'space'),

$$\mathcal{M} \cong \mathbb{R} \times \Sigma. \tag{4.38}$$

The topology of space–time is thus fixed. This may be a reasonable assumption in the classical theory, since topology change is usually connected with singularities or closed time-like curves. In the quantum theory, topology change may be a viable option, and its absence in the formalism could be a possible weakness of the canonical approach.[9] Nevertheless, the resulting quantum theory is general enough to cope with many of the interesting situations.

One therefore starts by performing a foliation of space–time into Cauchy surfaces Σ_t, with t denoting the global time function ('3+1 decomposition'). The corresponding vector field ('flow of time') is denoted by t^μ, obeying $t^\mu \nabla_\mu t = 1$. The relation between infinitesimally neighbouring hypersurfaces is the same as that shown in Fig. 3.1.[10] The space–time metric $g_{\mu\nu}$ induces a three-dimensional metric on each Σ_t according to

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu, \tag{4.39}$$

where n_μ denotes again the unit normal to Σ_t, with $n^\mu n_\mu = -1$.

[9] A more general formulation allowing topology change to occur in principle is the path-integral approach of Section 2.2.

[10] The vector field t^μ was called $\dot{X}^\mu$ in Fig. 3.1 and the relation (3.68).

This is in accordance with the earlier definition (3.69): multiplication of (4.39) by $X^\mu{}_{,a}X^\nu{}_{,b}$ and application of $X^\mu{}_{,a}n_\mu = 0$ leads to (3.69). In fact, $h_{\mu\nu}$ is a three-dimensional object only, since it acts as a projector on Σ_t; $h_{\mu\nu}n^\nu = 0$, $h_{\mu\nu}h^{\nu\rho} = h_\mu{}^\rho$. It is therefore really the three-dimensional metric, and we shall write h_{ab} for it below, since there is an isomorphism between tensor fields on $\mathcal{M}$ that are orthogonal to n^μ in each index and tensor fields on Σ_t.

As in (3.68), one can decompose t^μ into its components normal and tangential to Σ_t,

$$t^\mu = Nn^\mu + N^\mu, \tag{4.40}$$

where N is the lapse function and N^μ (called $N^a X^\nu{}_{,a}$ in (3.68)) is the shift vector. In fact, N^μ is a three-dimensional object and can be identified with N^a. The lapse function can be written as $N = -t^\mu n_\mu$, from which one can infer

$$N = \frac{1}{n^\mu \nabla_\mu t}. \tag{4.41}$$

As in the case of (3.22), one can interpret this expression as the ratio of the proper time (given by $t^\mu \nabla_\mu t = 1$) to the coordinate time $n^\mu \nabla_\mu t$. As in Section 3.3, the four-metric can be decomposed into spatial and temporal components,

$$g_{\mu\nu} = \begin{pmatrix} N_a N^a - N^2 & N_b \\ N_c & h_{ab} \end{pmatrix}. \tag{4.42}$$

Its inverse reads

$$g^{\mu\nu} = \begin{pmatrix} -\frac{1}{N^2} & \frac{N^b}{N^2} \\ \frac{N^c}{N^2} & h^{ab} - \frac{N^a N^b}{N^2} \end{pmatrix}. \tag{4.43}$$

Here, h^{ab} is the inverse of the three-metric (i.e. $h^{ab}h_{bc} = \delta^a_c$), and one recognizes that the spatial part of $g^{\mu\nu}$ is not identical to h^{ab} but contains an additional term involving the shift vector. The components of the normal vector can be found from the one-form $n_\mu\, \mathrm{d}x^\mu = -N\, \mathrm{d}t$ to be

$$n^\mu = g^{\mu\nu} n_\nu = \left(\frac{1}{N}, -\frac{\mathbf{N}}{N} \right). \tag{4.44}$$

The various hypersurfaces Σ_t can be identified by a diffeomorphism that is generated by the integral curves of t^μ. The globally hyperbolic space–time $(\mathcal{M}, g)$ can thus be interpreted as the time evolution of a Riemannian metric on a *fixed* manifold Σ, that is, as an evolution from $h_{ab}(t_0)$ to $h_{ab}(t)$. This suggests the use of the three-metric h_{ab} as the appropriate dynamical variable for the canonical formalism. Space–time then becomes nothing but a 'trajectory of spaces'. There is not even a need to assume from the beginning that Σ is embedded in some space–time; only *after* solving the equations of motion need we interpret $h_{ab}(t)$ as being brought about by 'wafting' Σ through a four-dimensional manifold $\mathcal{M}$ via a one-parameter family of embeddings.

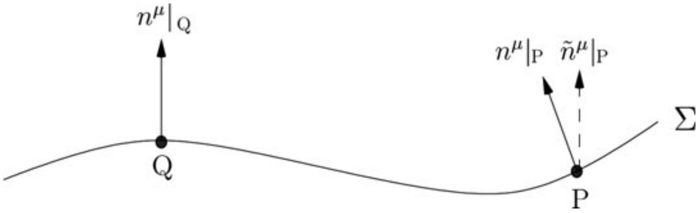

Fig. 4.2 Geometric interpretation of the extrinsic curvature.

In order to introduce the corresponding 'velocity' for h_{ab}, one can start by considering the following tensor field:[11]

$$K_{\mu\nu} = h_\mu^\rho \nabla_\rho n_\nu. \tag{4.45}$$

Since $K_{\mu\nu}n^\mu = 0 = K_{\mu\nu}n^\nu$, this tensor field is a purely spatial quantity and can be mapped to its spatial version K_{ab} (with indices raised and lowered by the three-metric). One can prove, using Frobenius' theorem for the hypersurface-orthogonal vector field n^μ, that this tensor field is symmetric, $K_{\mu\nu} = K_{\nu\mu}$.

Its geometric interpretation can be inferred from Fig. 4.2. Consider the normal vectors at two different points P and Q of a hypersurface. Let $\tilde{n}^\mu$ be the vector at P resulting from parallel-transporting n^μ along a geodesic from Q to P. The difference between n^μ and $\tilde{n}^\mu$ is a measure of the embedding curvature of Σ in $\mathcal{M}$ at P. One therefore recognizes that the tensor field (4.45) can be used to describe this embedding curvature, since it vanishes for $n^\mu = \tilde{n}^\mu$. One can also rewrite $K_{\mu\nu}$ in terms of a Lie derivative,

$$K_{\mu\nu} = \frac{1}{2}\mathcal{L}_\mathbf{n} h_{\mu\nu}, \tag{4.46}$$

where $\mathbf{n}$ denotes the unit normal vector field. Therefore, K_{ab} can be interpreted as the 'velocity' associated with h_{ab}. It is called the 'extrinsic curvature' or 'second fundamental form'. Its trace,

$$K := K_a{}^a = h^{ab}K_{ab} =: \theta, \tag{4.47}$$

can be interpreted as the 'expansion' of a geodesic congruence orthogonal to Σ.[12] In terms of lapse and shift, the extrinsic curvature can be written as

$$K_{ab} = \frac{1}{2N}\left(\dot{h}_{ab} - D_a N_b - D_b N_a\right), \tag{4.48}$$

and the two terms involving the spatial covariant derivative are together equivalent to $-\mathcal{L}_\mathbf{N} h_{ab}$. As we shall see in the next subsection, the components of the canonical momentum are obtained by a linear combination of the K_{ab}.

[11] Sometimes a different sign is used in this definition.

[12] For a Friedmann universe (cf. Section 8.1.2), K is three times the Hubble parameter.

4.2.2 Hamiltonian form of the Einstein–Hilbert action

One can now reformulate the Einstein–Hilbert action (1.1) in terms of the three-dimensional variables h_{ab} and K_{ab}. For this purpose, one needs the relationship between the four-dimensional and three-dimensional curvatures. This is given by the Gauss equation as generalized to higher dimensions,

$$ {}^{(3)}R_{\mu\nu\lambda}{}^{\rho} = h_{\mu}{}^{\mu'} h_{\nu}{}^{\nu'} h_{\lambda}{}^{\lambda'} h^{\rho}{}_{\rho'} R_{\mu'\nu'\lambda'}{}^{\rho'} - K_{\mu\lambda} K_{\nu}{}^{\rho} + K_{\nu\lambda} K_{\mu}{}^{\rho}, \tag{4.49} $$

and the generalized Codazzi equation,

$$ D_{\mu} K_{\nu\lambda} - D_{\nu} K_{\mu\lambda} = h_{\mu}{}^{\mu'} h_{\nu}{}^{\nu'} h_{\lambda}{}^{\lambda'} R_{\mu'\nu'\lambda'\rho} n^{\rho}. \tag{4.50} $$

Contraction of (4.50) with $g^{\mu\lambda}$ gives

$$ D_{\mu} K^{\mu}_{\ \nu} - D_{\nu} K = R_{\rho\lambda} n^{\lambda} h^{\rho}{}_{\nu}. \tag{4.51} $$

In the much simpler case of a two-dimensional hypersurface embedded in three-dimensional flat Euclidean space, (4.49) becomes the famous *theorema egregium* of Gauss after $R_{\mu\nu} = 0$ has been used for the embedding three-dimensional space (cf. the discussion by Kuchař 1993). In this case, the first term on the right-hand side is zero, and the *theorema* connects the only independent component of the two-dimensional Riemann tensor, ${}^{(2)}R_{2112}$, with the extrinsic curvature of the hypersurface,

$$ {}^{(2)}R_{2112} = K_{11} K_{22} - K_{21} K_{12} \equiv \det K_{ab}. \tag{4.52} $$

With the aid of the two principal curvatures κ_1 and κ_2, this can be written as

$$ {}^{(2)}R = 2\kappa_1 \kappa_2. \tag{4.53} $$

This gives the connection between intrinsic and extrinsic geometry, and its exact form holds because the embedding three-dimensional space is flat. If this embedding space has a Lorentzian signature, one gets instead

$$ {}^{(2)}R = -2\kappa_1 \kappa_2. \tag{4.54} $$

Kuchař (1993) expresses this in the form that the 'law of the instant' (because the hypersurface refers to $t = $ constant) implies the 'dynamical law' (expressing the flatness of the whole embedding space–time). For this conclusion to be true, the condition (4.54) must hold for *all* hypersurfaces.

In 3+1 dimensions, the situation is more complicated. Examining the vacuum Einstein equations,[13] $G_{\mu\nu} = 0$, one finds for the 'space–time component' G_{i0}

$$ 0 = h^{\mu}{}_{\rho} G_{\mu\nu} n^{\nu} = h^{\mu}{}_{\rho} R_{\mu\nu} n^{\nu}, $$

which can be rewritten with the help of (4.51) as

$$ D_b K^{b}{}_{a} - D_a K = 0. \tag{4.55} $$

[13] The cosmological constant is neglected here for simplicity.

For the 'time–time component', one has

$$0 = G_{\mu\nu}n^\mu n^\nu = R_{\mu\nu}n^\mu n^\nu + \frac{R}{2}. \tag{4.56}$$

From (4.49), one finds upon contraction of indices

$${}^{(3)}R + K_\mu{}^\mu K_\nu{}^\nu - K_{\mu\nu}K^{\mu\nu} = h^{\mu\mu'}h_\nu{}^{\nu'}h_\mu{}^{\lambda'}h^\nu{}_{\rho'}R_{\mu'\nu'\lambda'}{}^{\rho'}.$$

The right-hand side is equal to

$$R + 2R_{\mu\nu}n^\mu n^\nu = 2G_{\mu\nu}n^\mu n^\nu,$$

and so the 'time–time component' of Einstein's equations reads

$$K^2 - K_{ab}K^{ab} + {}^{(3)}R = 0. \tag{4.57}$$

This is the (3+1)-dimensional version of the *theorema egregium*. Both (4.55) and (4.57) are *constraints*—they only contain first-order time derivatives. The space–space components of the Einstein equations are dynamical.

The constraints play a crucial role in the initial-value formulation of classical GR; see, for example, Choquet-Bruhat and York (1980) for a review. There, one can specify initial data (h_{ab}, K_{cd}) on Σ, where h_{ab} and K_{cd} satisfy the constraints (4.55) and (4.57). One can then prove that there exists one globally hyperbolic space–time obeying Einstein's equations (i.e. a unique solution for the four-metric up to diffeomorphisms), which has a Cauchy surface on which the induced metric and the extrinsic curvature are just h_{ab} and K_{cd}, respectively.

In electrodynamics, for comparison, one has to specify $\mathbf{A}$ and $\mathbf{E}$ on Σ satisfying the constraint (Gauss's law (4.30)) $\nabla\mathbf{E} = 0$. One then gets in space–time a solution of Maxwell's equations that is unique up to gauge transformation. The important point is that the space–time is fixed in Maxwell's theory, whereas in the gravitational case it is part of the solution.

That the dynamical laws follow from the laws of the instant can be inferred from the validity of the following 'interconnection theorems' (cf. Kuchař (1981)):

1. If the constraints are valid on an initial hypersurface and if the dynamical evolution equations $G_{ab} = 0$ (pure spatial components of the vacuum Einstein equations) on space–time hold, the constraints hold on *every* hypersurface. Together, one then has all ten Einstein equations.
2. If the constraints hold on every hypersurface, the equations $G_{ab} = 0$ hold on space–time.

Similar properties hold in Maxwell's theory; see Giulini and Kiefer (2007). In the presence of non-gravitational fields, $\nabla_\mu T^{\mu\nu} = 0$ is needed as an integrability condition (analogously to $\partial_\mu j^\mu = 0$ for Maxwell's equations).

In order to reformulate the Einstein–Hilbert action (1.1), one has to express the volume element and the Ricci scalar in terms of h_{ab} and K_{cd}. For the volume element, one finds

$$\sqrt{-g} = N\sqrt{h}. \tag{4.58}$$

This can be seen as follows. Defining the three-dimensional volume element as (see e.g. Wald 1984)

$$^{(3)}e_{\mu\nu\lambda} = e_{\rho\mu\nu\lambda}t^\rho,$$

with t^ρ according to (4.40) and $e_{\rho\mu\nu\lambda}$ denoting the time-independent four-dimensional volume element, one has, by using $\epsilon_{\rho\mu\nu\lambda} = \sqrt{-g}\, e_{\rho\mu\nu\lambda}$,

$$\epsilon_{\rho\mu\nu\lambda}t^\rho = \sqrt{-g}\, e_{\mu\nu\lambda} = \sqrt{-\frac{g}{h}}\epsilon_{\mu\nu\lambda},$$

from which (4.58) follows after using (4.40) and taking purely spatial components. Equation (4.58) can also be found from (4.42).

We shall now assume in the following that Σ is compact without boundary; the boundary terms for the non-compact case will be discussed separately in Section 4.2.4. In order to rewrite the curvature scalar, we first use (4.56) in the following form:

$$R = {}^{(3)}R + K^2 - K_{ab}K^{ab} - 2R_{\mu\nu}n^\mu n^\nu. \tag{4.59}$$

Using the definition of the Riemann tensor in terms of second covariant derivatives,

$$R^\rho{}_{\mu\rho\nu}n^\mu = \nabla_\rho\nabla_\nu n^\rho - \nabla_\nu\nabla_\rho n^\rho,$$

the second term on the right-hand side can be written as

$$\begin{aligned} -2R_{\mu\nu}n^\mu n^\nu &= 2(\nabla_\rho n^\nu)(\nabla_\nu n^\rho) - 2\nabla_\rho(n^\nu\nabla_\nu n^\rho) \\ &\quad -2(\nabla_\nu n^\nu)(\nabla_\rho n^\rho) + 2\nabla_\nu(n^\nu\nabla_\rho n^\rho). \end{aligned} \tag{4.60}$$

The second and the fourth term are total divergences. They can thus be cast into surface terms at the temporal boundaries. The first surface term yields $-2(n^\nu\nabla_\nu n^\rho)$ $n_\rho = 0$, while the second one gives $2\nabla_\mu n^\mu = -2K$ (recall (4.45)). The two remaining terms in (4.60) can be written as $2K_{ab}K^{ab}$ and $-2K^2$, respectively. Inspecting the Einstein–Hilbert action (1.1), one recognizes that the temporal surface term is cancelled, and that the action now reads

$$\begin{aligned} 16\pi G\, S_{\text{EH}} &= \int_{\mathcal{M}} \mathrm{d}t\, \mathrm{d}^3x\ N\sqrt{h}(K_{ab}K^{ab} - K^2 + {}^{(3)}R - 2\Lambda) \\ &\equiv \int_{\mathcal{M}} \mathrm{d}t\, \mathrm{d}^3x\ N\left(G^{abcd}K_{ab}K_{cd} + \sqrt{h}[{}^{(3)}R - 2\Lambda]\right), \end{aligned} \tag{4.61}$$

where, in the second line, DeWitt's metric (4.25) has been introduced. The action (4.61) is also called the 'ADM action' in recognition of the work by Arnowitt, Deser, and Misner; see Arnowitt *et al.* (1962). It has the classic form of kinetic energy minus potential energy, since the extrinsic curvature contains the 'velocities' $\dot{h}_{ab}$; see (4.48). Writing

$$S_{\mathrm{EH}} =: \int_{\mathcal{M}} \mathrm{d}t\, \mathrm{d}^3x\ \mathcal{L}^{\mathrm{g}},$$

one gets the following expressions for the canonical momenta. First,

$$p_N := \frac{\partial \mathcal{L}^{\mathrm{g}}}{\partial \dot{N}} = 0\ , \quad p_a^{\mathrm{g}} := \frac{\partial \mathcal{L}^{\mathrm{g}}}{\partial \dot{N}^a} = 0. \tag{4.62}$$

Because the lapse function and shift vector are only Lagrange multipliers (like A^0 in electrodynamics), these are constraints; because they do not involve the dynamical equations, they are primary constraints; cf. Section 3.5. Second,

$$p^{ab} := \frac{\partial \mathcal{L}^{\mathrm{g}}}{\partial \dot{h}_{ab}} = \frac{1}{16\pi G} G^{abcd} K_{cd} = \frac{\sqrt{h}}{16\pi G}\left(K^{ab} - K h^{ab}\right). \tag{4.63}$$

Note that the gravitational constant G appears here explicitly, although no coupling to matter is involved. This is the reason why it will appear in vacuum quantum gravity; see Section 5.2. One therefore has the Poisson-bracket relation[14]

$$\{h_{ab}(x), p^{cd}(y)\} = \delta^c_{(a}\delta^d_{b)}\delta(x,y). \tag{4.64}$$

Recalling (4.48) and taking the trace of (4.63), one can express the velocities in terms of the momenta:

$$\dot{h}_{ab} = \frac{32\pi G N}{\sqrt{h}}\left(p_{ab} - \frac{1}{2} p h_{ab}\right) + D_a N_b + D_b N_a, \tag{4.65}$$

where $p := p^{ab}h_{ab}$. One can now calculate the canonical Hamiltonian density

$$\mathcal{H}^{\mathrm{g}} = p^{ab}\dot{h}_{ab} - \mathcal{L}^{\mathrm{g}},$$

for which one gets the expression[15]

$$\mathcal{H}^{\mathrm{g}} = 16\pi G N G_{abcd} p^{ab} p^{cd} - N\frac{\sqrt{h}({}^{(3)}R - 2\Lambda)}{16\pi G} - 2N_b(D_a p^{ab}). \tag{4.66}$$

The full Hamiltonian is found by integration:

$$H^{\mathrm{g}} := \int \mathrm{d}^3x\ \mathcal{H}^{\mathrm{g}} := \int \mathrm{d}^3x\ (N\mathcal{H}^{\mathrm{g}}_{\perp} + N^a\mathcal{H}^{\mathrm{g}}_a). \tag{4.67}$$

The action (4.61) can be written in the form

$$16\pi G\ S_{\mathrm{EH}} = \int \mathrm{d}t\, \mathrm{d}^3x\ \left(p^{ab}\dot{h}_{ab} - N\mathcal{H}^{\mathrm{g}}_{\perp} - N^a\mathcal{H}^{\mathrm{g}}_a\right). \tag{4.68}$$

[14]This is formal at this stage, since it does not take into account the fact that $\sqrt{h} > 0$.

[15]This holds modulo a total divergence, which does not contribute in the integral because Σ is compact.

Variation with respect to the Lagrange multipliers N and N^a yields the constraints[16]

$$\mathcal{H}_\perp^{\mathrm{g}} = 16\pi G\, G_{abcd} p^{ab} p^{cd} - \frac{\sqrt{h}}{16\pi G}({}^{(3)}R - 2\Lambda) \approx 0 \ , \tag{4.69}$$

$$\mathcal{H}_a^{\mathrm{g}} = -2D_b p_a{}^b \approx 0. \tag{4.70}$$

In fact, the constraint (4.69) is equivalent to (4.57), and (4.70) is equivalent to (4.55)—they are called the Hamiltonian constraint and the diffeomorphism (or momentum) constraints, respectively. From its structure, (4.69) has some similarity to the constraint for the relativistic particle (3.18), while (4.70) is similar to (4.30).

It can now be seen explicitly that these constraints are equivalent to the results from the 'seventh route to geometrodynamics'; see (4.21) and (4.22). The total Hamiltonian is thus constrained to vanish, a result that is in accordance with our general discussion of reparametrization invariance in Section 3.1. In the case of non-compact space, boundary terms are present in the Hamiltonian; see Section 4.2.4. It is also obvious from the 'seventh route' that the constraints (4.69) and (4.70) obey the algebra of constraints given in (3.84)–(3.86). Since this algebra closes on the constraints themselves, the constraints are of first class in the terminology of Section 3.5.

In addition to the constraints, one has the six dynamical equations, the Hamiltonian equations of motion. The first half, $\dot{h}_{ab} = \{h_{ab}, H^{\mathrm{g}}\}$, just gives (4.65). The second half, $\dot{p}^{ab} = \{p^{ab}, H^{\mathrm{g}}\}$, yields a lengthy expression (see e.g. Wald 1984) that is not needed for canonical quantization; the reason is that, as we shall see in Chapter 5, the whole of the information about the quantum theory is in the constraints. The expression is, of course, needed for applications of the classical canonical formalism such as gravitational-wave emission from compact binary objects.

If non-gravitational fields are coupled, the constraints acquire extra terms. In (4.56), one has to use the fact that

$$2G_{\mu\nu}n^\mu n^\nu = 16\pi G T_{\mu\nu} n^\mu n^\nu =: 16\pi G\rho.$$

Instead of (4.69), one now has the following expression for the Hamiltonian constraint:

$$\mathcal{H}_\perp = 16\pi G\, G_{abcd} p^{ab} p^{cd} - \frac{\sqrt{h}}{16\pi G}({}^{(3)}R - 2\Lambda) + \sqrt{h}\rho \approx 0. \tag{4.71}$$

Similarly, instead of (4.70), one has the following expression for the diffeomorphism constraints:

$$\mathcal{H}_a = -2D_b p_a{}^b + \sqrt{h} J_a \approx 0, \tag{4.72}$$

where $J_a := h_a{}^\mu T_{\mu\nu} n^\nu$ is the 'Poynting vector'. Consider as special examples the cases of a scalar field and the electromagnetic field. With the Lagrange density

$$\mathcal{L} = \sqrt{-g}\left(-\tfrac{1}{2}\, g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} - \tfrac{1}{2}\, m^2\phi^2\right) \tag{4.73}$$

for the scalar field, one finds for its Hamiltonian

[16] These also follow from the preservation of the primary constraints, $\{p_N, H^{\mathrm{g}}\} = 0 = \{p_a^{\mathrm{g}}, H^{\mathrm{g}}\}$.

$$H_\phi = \int \mathrm{d}^3x\ N\left(\frac{p_\phi^2}{2\sqrt{h}} + \frac{\sqrt{h}}{2}h^{ab}\phi_{,a}\phi_{,b} + \frac{1}{2}\sqrt{h}m^2\phi^2\right) + \int \mathrm{d}^3x\ N^a p_\phi \phi_{,a}. \tag{4.74}$$

The term in parentheses in the first integral has to be added to $\mathcal{H}_\perp^{\mathrm{g}}$, while the term $p_\phi\phi_{,a}$ (which we have already encountered in (4.14)) must be added to $\mathcal{H}_a^{\mathrm{g}}$. For the electromagnetic field, starting from

$$\mathcal{L} = -\tfrac{1}{4}\sqrt{-g}g^{\mu\rho}g^{\nu\sigma}F_{\mu\nu}F_{\rho\sigma}, \tag{4.75}$$

one gets for the Hamiltonian

$$H_{\mathrm{EM}} = \int \mathrm{d}^3x\ N\left(\frac{h_{ab}}{2\sqrt{h}}p^a p^b + \frac{\sqrt{h}}{2}h^{ab}B_aB_b\right) + \int \mathrm{d}^3x\ N^a\left(\partial_a A_b - \partial_b A_a\right)p^b - \int \mathrm{d}^3x\ A_0\partial_a p^a, \tag{4.76}$$

where $p^a = \partial\mathcal{L}/\partial\dot{A}_a$ is the electric field and $B_a = (1/2)\epsilon_{abc}F^{bc}$ is the magnetic field. Note that the term in parentheses in the second integral is just the $\bar{\mathcal{H}}_a$ from (4.28), and variation with respect to the Lagrange multiplier A_0 yields Gauss's law (4.30). If fermionic degrees of freedom are present, one must perform the 3+1 decomposition with respect to the vierbein instead of the four-metric; see for example, Ashtekar (1991).

The classical canonical formalism for the gravitational field as discussed up to now was pioneered by Rosenfeld (1930*b*), Dirac (1958*b*, 1959), Bergmann's group (Bergmann 1989, 1992), and 'ADM' (Arnowitt *et al.* 1962); cf. also Section 3.5.

We remark, finally, that the canonical quantization of higher-derivative theories such as R^2-gravity can also be performed (cf. Boulware 1984). The formalism is then more complicated, since one has to introduce additional configuration-space variables and momenta.

4.2.3 Discussion of the constraints

The presence of the constraints derived in the last subsection means that only part of the variables constitute physical degrees of freedom (cf. Section 3.1.2). How may one count them? The three-metric $h_{ab}(x)$ is characterized by six numbers per space point (often symbolically denoted as $6\times\infty^3$). The diffeomorphism constraints (4.70) generate coordinate transformations on three-space. These are characterized by three numbers, so $6-3=3$ numbers per point remain. The constraint (4.69) corresponds to one variable per space point describing the location of Σ in space–time (since Σ changes under normal deformations). In a sense, this one variable therefore corresponds to 'time', and $2\times\infty^3$ degrees of freedom remain. Baierlein *et al.* (1962) have interpreted this as the 'three-geometry carrying information about time'.

The gravitational field thus seems to be characterized by $2\times\infty^3$ intrinsic degrees of freedom. This is consistent with the corresponding number found in linear quantum gravity—the two spin-2 states of the graviton (Section 2.1). One can, alternatively, perform the following counting in phase space. The canonical variables $(h_{ab}(x), p^{cd}(y))$ are $12 \times \infty^3$ variables. Due to the presence of the four constraints in phase space, $4 \times \infty^3$ variables have to be subtracted. The remaining $8 \times \infty^3$ variables define the constraint hypersurface Γ_c. Since the constraints are of first class, they generate a four-parameter set of redundancy or 'gauge' transformations on Γ_c (see Section 3.5), and $4 \times \infty^3$ degrees of freedom must be subtracted in order to 'fix the gauge'. The remaining $4\times\infty^3$ variables define the reduced phase space Γ_r and correspond to $2\times\infty^3$ degrees of freedom in configuration space—in accordance with the counting above. It must be emphasized that this counting always holds modulo a *finite* number of degrees of freedom—an example is gravity in 2+1 dimensions (Section 8.1.3).

Does a three-dimensional geometry indeed contain information about time? Consider a situation in non-gravitational physics, for example electrodynamics. There, the specification of, say, the magnetic field on two hypersurfaces does not suffice to determine the field everywhere. In addition, the time parameters of the two surfaces must be specified for an appropriate boundary-value problem. In contrast to the gravitational case, the background space–time is fixed here (i.e. it is non-dynamical). As we have seen in Section 3.4, two configurations (e.g. of a clock) in classical mechanics do not suffice to determine the motion—one needs in addition the two times of the clock configurations or its speed.

The situation in the gravitational case is related to the 'sandwich conjecture'. This conjecture states that two three-geometries do (in the generic case) determine the temporal separation (the proper times) along each time-like worldline connecting them in the resulting space–time. Whereas not much is known about the finite version of this conjecture, results are available for the infinitesimal case. In this 'thin-sandwich conjecture', one specifies on one hypersurface the three-metric h_{ab} and its 'time derivative' $\partial h_{ab}/\partial t$—the latter is only required up to a numerical factor, since the 'speed' itself is meaningless; only the 'direction' in configuration space is of significance. The thin-sandwich conjecture holds if one can determine the lapse and shift from these initial conditions and the constraints.[17] It has been shown that this can be done locally for 'generic' situations; see Bartnik and Fodor (1993) for pure gravity and Giulini (1999) for gravity plus matter.[18]

The 'temporal' degree of freedom of the three-geometry cannot in general be separated from other variables; that is, all three degrees of freedom contained in h_{ab} (after the diffeomorphism constraints have been considered) should be interpreted as physical variables, and be treated on an equal footing. In the special case of linear gravity (Section 2.1), a background structure is present. This enables one to separate a distinguished time and to regard the remaining variables, the two degrees of freedom of the graviton, as the only physical variables. The identification of one variable contained in the $h_{ab}(x)$ as 'time' thus seems only possible in situations where the hypersurface is already embedded in a space–time satisfying Einstein's equations (which will cer-

[17]This boundary-value problem must be distinguished from the one above where h_{ab} and p^{cd} are specified on Σ and a space–time can be chosen after the lapse and shift are *freely* chosen.

[18]The condition is that the initial speed must have at each space point a negative square with respect to the DeWitt metric.

tainly not be the case in quantum gravity; see Chapter 5). Moreover, Torre (1993*a*) has shown that GR cannot be equivalent to a deparametrized theory in the sense of Section 3.3; that is, no distinguished time variable can be retrieved. This will have some relevance for the discussion of time in quantum gravity; see Section 5.2.1.

Starting from the ADM action (4.61), one may find an alternative formulation by first varying the action with respect to N and then inserting the ensuing solution back into it. This corresponds to the solution of the Hamiltonian constraint. Writing for the DeWitt metric

$$G^{abcd} =: \sqrt{h}\mathcal{G}^{abcd},$$

a variation with respect to N yields (for $\Lambda = 0$)

$$N = \frac{1}{2}\sqrt{\frac{\mathcal{G}^{abcd}(\dot{h}_{ab} - 2D_{(a}N_{b)})(\dot{h}_{cd} - 2D_{(c}N_{d)})}{{}^{(3)}R}}. \tag{4.77}$$

Reinserting (4.77) into (4.61), one finds the 'Baierlein–Sharp–Wheeler' (BSW) form of the action,

$$16\pi G S_{\mathrm{BSW}} = \int \mathrm{d}t\, \mathrm{d}^3x\, \sqrt{h}\sqrt{{}^{(3)}R\mathcal{G}^{abcd}(\dot{h}_{ab} - 2D_{(a}N_{b)})(\dot{h}_{cd} - 2D_{(c}N_{d)})}. \tag{4.78}$$

A justification of S_{BSW} from a 'Machian' viewpoint can be found in Barbour *et al.* (2002). This action resembles the Jacobi-type action (3.99), but is much more sophisticated. Concerning the lapse and shift, now only the shift functions N^a have to be varied. If a unique solution existed, one could employ the thin-sandwich conjecture; that is, one could construct the space–time out of initial data h_{ab} and $\dot{h}_{ab}$. This procedure corresponds to the 'horizontal stacking' (also called 'best matching' or the 'G-frame method') mentioned in Section 3.4. In such an approach, GR can be derived without any prior assumptions of a space–time nature such as the validity of the principle of path independence described in Section 4.1. It is not necessary to require the constraints to close in the specific manner of the algebra of surface deformations (3.84)–(3.86). Mere closure in any fashion is sufficient to accommodate GR, universality of the lightcone, and gauge theory (Barbour *et al.* 2002; cf. also Anderson 2007*a*).

In this connection, Barbour (1994) argues that the Hamiltonian constraint does not further restrict the number of physical variables (i.e. it does not restrict them from $3 \times \infty^3$ to $2 \times \infty^3$), but is an identity that reflects the fact that only the *direction* of the initial velocity matters, not its absolute value. In the model of Section 3.4, one has for the canonical momentum the expression

$$\mathbf{p}_k = \sqrt{-V}\frac{m_k \dot{\mathbf{x}}_k}{\sqrt{\frac{1}{2}\sum_k m_k \dot{\mathbf{x}}_k^2}}, \tag{4.79}$$

leading to the constraint (3.102). The second factor in (4.79) resembles the expression for 'direction cosines',[19] their usual relation—the sum of their squares is equal to

[19] Direction cosines are the cosines of the angles between a vector and the cartesian coordinate axes.

one—being equivalent to (3.102). In other words, this factor remains unchanged under a scale transformation of the velocities.

Here one finds the following from (4.78) for the momentum:[20]

$$p^{ab}_{\mathrm{BSW}} = \frac{\sqrt{h}}{16\pi G}\sqrt{\frac{{}^{(3)}R}{T}}\mathcal{G}^{abcd}(\dot{h}_{cd} - 2D_{(c}N_{d)}), \tag{4.80}$$

where the 'kinetic term' is given by

$$T = \mathcal{G}^{abcd}(\dot{h}_{ab} - 2D_{(a}N_{b)})(\dot{h}_{cd} - 2D_{(c}N_{d)}). \tag{4.81}$$

Similarly to (4.79), these momenta resemble direction cosines, which, due to the infinite dimensions of configuration space, are 'local' direction cosines.

Since the Hamiltonian constraint (4.69) and the momentum constraints (4.70) are part of the total Hamiltonian, the following conceptual problem arises (see Pons *et al.* 2009 and the references therein). On the one hand, first-class constraints generate redundancy transformations. On the other hand, the Hamiltonian is supposed to generate time evolution. How can a redundancy transformation be equivalent to a time evolution? In order to answer this question, let us recall from the general discussion in Section 3.5 that the generator of a redundancy transformation is of the form (3.118), where the various constraints of the theory are multiplied by arbitrary functions ξ and their time derivatives. The total Hamiltonian in the sense of the Dirac–Bergmann algorithm reads

$$H_{\mathrm{T}} = N^{\mu}\mathcal{H}_{\mu} + \lambda^{\mu}P_{\mu}, \tag{4.82}$$

where N^{μ} comprises N and N^a, and $\mathcal{H}_{\mu}$ comprises $\mathcal{H}^{\mathrm{g}}_{\perp}$ and $\mathcal{H}^{\mathrm{g}}_{a}$; P_{μ} stands for the collection of the primary constraints, p_N and p^{g}_{a}. The sums in (4.82) include, for convenience, a spatial integration. The redundancy generator (3.118) turns out to have the following explicit form:

$$G_{\xi}(t) = (\mathcal{H}_{\mu} + N^{\rho}U^{\nu}_{\mu\rho}P_{\nu})\xi^{\mu} + P_{\mu}\dot{\xi}^{\mu}. \tag{4.83}$$

Here, the $U^{\nu}_{\mu\rho}$ summarize the structure functions that occur in the constraint algebra (3.84)–(3.86). This generator transforms an entire solution (given by N^{μ} and h_{ab}) into another (equivalent) entire solution (given by $\tilde{N}^{\mu}$ and $\tilde{h}_{ab}$). We emphasize that the generator contains both the primary and the secondary (first-class) constraints; both are needed to generate a redundancy ('gauge') transformation. In contrast to G_{ξ}, H_{T} acts within a single solution and can thus be interpreted as the generator of a rigid (many-fingered) time evolution starting from initial conditions. First-class constraints in the Hamiltonian can thus well coexist with time translations. The Hamiltonian in GR can generate physical motion (Barbour and Foster 2008).

An important concept in this context is the notion of an observable. An observable was defined in Section 3.5 as a variable that weakly commutes with the constraints. Here, an observable $\mathcal{O}$ should then satisfy

$$\{\mathcal{O}, \mathcal{H}^{\mathrm{g}}_{a}\} \approx 0\ , \tag{4.84}$$

$$\{\mathcal{O}, \mathcal{H}^{\mathrm{g}}_{\perp}\} \approx 0. \tag{4.85}$$

[20]This is of course equal to p^{ab} after the identification (4.77) has been made for the lapse.

(There is also a weak commutativity with the primary constraints.) While the first condition is intuitively clear (observables should not depend on the chosen coordinates of Σ), the situation is less clear for the second condition. This has to do with the above-mentioned double role of the Hamiltonian as a generator of time translations and as a constraint. Kuchař (1993) called all quantities that obey (4.84) observables, independent of whether they obey (4.85) or not. He called variables that obey *in addition* (4.85) 'perennials'. The reason for this is the association of perennials with constants of motion.

However, weak commutativity with the Hamiltonian does not necessarily lead to constants of motion. The reason is that there may be an explicit time dependence in the observables. Such an explicit dependence can arise from imposing explicit coordinate conditions ('gauge fixing'); see Pons *et al.* (2009). Such a gauge fixing is performed by choosing a set of four independent scalar fields X^μ that obey $x^\mu - X^\mu(x) = 0$, which brings the coordinates (and therefore the time dependence) explicitly into play. The idea is that this set is realized in nature by physical objects such as the GPS satellites or by geometrical scalars such as the four scalars of the Weyl tensor. The latter case gives a particular example in which time-dependent observables are present (Bergmann and Komar 1960). Observables that are constants of motion can be constructed from the time-dependent observables by a well-defined algorithm (Pons *et al.* 2009). They bear a close resemblance to the 'evolving constants of motion' discussed by Rovelli (1991*b*).

To find explicit expressions for observables in GR is, at least for the case of a compact three-space, an extremely difficult task, and no single example is known so far. Torre (1993*b*) has shown that there can be no observables satisfying (4.84) and (4.85) for the vacuum gravitational field (in a closed universe) built as spatial integrals of local functions of Cauchy data; observables, if they exist, must thus be highly non-local quantities.

We have already seen in Section 4.1 that the transformations generated by the constraints (4.69) and (4.70) are different from the original space–time diffeomorphisms of GR. The formal reason is that $\mathcal{H}_\perp^{\mathrm{g}}$ is non-linear in the momenta, so the transformations in the phase space Γ spanned by (h_{ab}, p^{cd}) cannot be reduced to space–time transformations. What, then, is the relation between the two types of transformations? Let $(\mathcal{M}, g)$ be a globally hyperbolic space–time. We shall denote by Riem $\mathcal{M}$ the space of all four-dimensional (pseudo-)Riemannian metrics on $\mathcal{M}$. Since the group of space–time diffeomorphisms, Diff $\mathcal{M}$, does not act transitively, there exist non-trivial orbits in Riem $\mathcal{M}$. One can make a projection down to the space of all four-geometries, Riem $\mathcal{M}$/Diff $\mathcal{M}$. By considering a particular section,

$$\sigma : \text{Riem } \mathcal{M}/\text{Diff } \mathcal{M} \mapsto \text{Riem } \mathcal{M}, \tag{4.86}$$

one can choose a particular representative metric on $\mathcal{M}$ for each geometry. In this way, one can define formal points of the 'background manifold' $\mathcal{M}$, which a priori have no meaning (in GR, points cannot be disentangled from the metric fields). The map between different sections is *not* a single diffeomorphism, but a more complicated transformation; in phase space, it corresponds to an element of the 'Bergmann–Komar group' (Bergmann and Komar 1972), whose generator is given by (4.83), cf. Pons *et al.* (2009). The Bergmann–Komar group is the maximal subgroup of the (field-dependent) diffeomorphism group that can be realized in phase space. It is this group that is the imprint of the space–time diffeomorphism group in the canonical formalism.

Concerning the space Riem $\mathcal{M}$ of four-metrics, Hájíček and Kijowski (2000) have shown (see also Hájíček and Kiefer 2001*a* and Section 7.2) that there exists a map from

$$\text{Riem } \mathcal{M}/\text{Diff } \mathcal{M} \times \text{Emb}(\Sigma, \mathcal{M}),$$

where $\text{Emb}(\Sigma, \mathcal{M})$ denotes the space of all embeddings of Σ into $\mathcal{M}$, into the phase space Γ but *excluding* points where the constructed space–times admit an isometry. Therefore, the identification between space–time diffeomorphisms and the transformations in phase space proceeds via whole 'histories'. The necessary exclusion of points representing Cauchy data for space–times with Killing vectors from Σ is one of the reasons why GR cannot be equivalent to a deparametrized theory (Torre 1993*a*).

One interesting limit for the Hamiltonian constraint (4.69) is the 'strong-coupling limit' defined by formally setting $G \to \infty$. This is the limit opposite to the weak-coupling expansion of Chapter 2. It also corresponds formally to the limit $c \to 0$, that is, the limit opposite to the Galilean case of an infinite speed of light. This can be seen by noting that the constant in front of the potential term in (4.69) in fact reads $c^4/16\pi G$. Therefore, in this limit, the lightcones effectively collapse to the axes $x =$ constant; different spatial points decouple because all spatial-derivative terms present in ${}^{(3)}R$ disappear. One can show that this situation corresponds to having a 'Kasner universe' at each space point; see, for example, Pilati (1982, 1983) for details. Since the potential term also carries the signature σ, this limit corresponds formally to $\sigma = 0$; that is, the Poisson bracket between Hamiltonian constraints (3.84) becomes zero. The decoupling of space points can also be recognized in the BKL oscillations that occur when one approaches the cosmological singularity; cf. Belinskii *et al.* (1982).

4.2.4 The case of open spaces

Up to now, we have neglected the presence of possible spatial boundary terms in the Hamiltonian. In this subsection, we shall briefly discuss the necessary modifications for the case of open spaces (where 'open' means 'asymptotically flat'). The necessary details can be found in Regge and Teitelboim (1974) and Beig and Ó Murchadha (1987).

Variation of the full Hamiltonian H^{g} with respect to the canonical variables h_{ab} and p^{cd} yields

$$\delta H^{\mathrm{g}} = \int \mathrm{d}^3x \ \left(A^{ab}\delta h_{ab} + B_{ab}\delta p^{ab}\right) - \delta C, \tag{4.87}$$

where δC denotes surface terms. Because H^{g} must be a differentiable function with respect to h_{ab} and p^{cd} (otherwise Hamilton's equations of motion would not make sense), δC must be *cancelled* by introducing explicit surface terms into H^{g}. For the derivation of such surface terms, one must impose fall-off conditions on the canonical variables. For the three-metric they read

$$h_{ab} \overset{r\to\infty}{\sim} \delta_{ab} + \mathcal{O}\left(\frac{1}{r}\right) \ , \quad h_{ab,c} \overset{r\to\infty}{\sim} \mathcal{O}\left(\frac{1}{r^2}\right), \tag{4.88}$$

and analogously for the momenta. The lapse and shift, if again combined into the four-vector N^μ, are required to obey

$$N^\mu \overset{r\to\infty}{\sim} \alpha^\mu + \beta^\mu_a x^a, \tag{4.89}$$

where α^μ describe space–time translations, $\beta_{ab} = -\beta_{ba}$ spatial rotations, and $\beta_a^\perp$ boosts. Together, they form the Poincaré group of Minkowski space–time, which is a symmetry in the asymptotic sense. The procedure mentioned above then leads to the following expression for the total Hamiltonian:

$$H^{\mathrm{g}} = \int \mathrm{d}^3x \; (N\mathcal{H}^{\mathrm{g}}_\perp + N^a\mathcal{H}^{\mathrm{g}}_a) + \alpha E_{\mathrm{ADM}} - \alpha^a P_a + \tfrac{1}{2}\, \beta_{\mu\nu} J^{\mu\nu}, \tag{4.90}$$

where E_{ADM} (also called the 'ADM' energy; see Arnowitt *et al.* 1962), P_a, and $J^{\mu\nu}$ are the total energy, the total momentum, and the total angular momentum plus the generators of boosts, respectively.[21] Together they form the generators of the Poincaré group at infinity. They obey the standard commutation relations (2.34)–(2.36). For the ADM energy, in particular, one finds the expression

$$E_{\mathrm{ADM}} = \frac{1}{16\pi G} \oint_{r\to\infty} \mathrm{d}^2\sigma_a\, (h_{ab,b} - h_{bb,a}). \tag{4.91}$$

Note that the total energy is defined by a surface integral over a sphere for $r \to \infty$ and *not* by a volume integral. One can prove that $E_{\mathrm{ADM}} \geq 0$.

The integral in (4.90) is the same integral as in (4.67). Because of the constraints (4.69) and (4.70), H^{g} is numerically equal to the surface terms. For vanishing asymptotic shift and a lapse equal to one, H^{g} is just given by the ADM energy. We emphasize that the asymptotic Poincaré transformations must not be interpreted as gauge transformations (otherwise E_{ADM}, P^a, and $J^{\mu\nu}$ would be constrained to vanish) but as proper physical symmetries; see the remarks in the following subsection.

4.2.5 Structure of configuration space

An important preparation for the quantum theory is an investigation into the structure of the configuration space because this will be the space on which the wave functional will be defined.

We have seen that the canonical formalism deals with the set of all three-metrics on a given manifold Σ. We call this space Riem Σ (not to be confused with Riem $\mathcal{M}$ considered above). As the configuration space of the theory, we want to address the quotient space in which all metrics corresponding to the same three-geometry are identified. Following Wheeler (1968), we call this space 'superspace'. It is defined by

$$\mathcal{S}(\Sigma) := \mathrm{Riem}\ \Sigma/\mathrm{Diff}\ \Sigma. \tag{4.92}$$

By going to superspace, the momentum constraints (4.70) are automatically fulfilled. Whereas Riem Σ has a simple topological structure (it is a cone in the vector space of all symmetric second-rank tensor fields), the topological structure of $\mathcal{S}(\Sigma)$ is very

[21] They are observables in the sense of (4.84) and (4.85).

complicated because it inherits (through Diff Σ) some of the topological information contained in Σ.

In general, Diff Σ can be divided into a 'symmetry part' and a 'redundancy part' (Giulini 1995*a*, 2009). Symmetries arise typically in the case of asymptotically flat spaces (Section 4.2.4). They describe, for example, rotations with respect to the remaining part of the universe ('fixed stars'). Since they have physical significance, they should not be factored out, and Diff Σ is then understood to contain only the 'true' diffeomorphisms (redundancies). In the closed case (relevant in particular for cosmology), only the redundancy part is present.

For closed Σ, $\mathcal{S}(\Sigma)$ has a non-trivial singularity structure due to the occurrence of metrics with isometries (Fischer 1970); at such singular points, superspace is not a manifold (a situation like that, for example, at the tip of a cone). A proposal to avoid such singularities by employing a 'resolution space' $\mathcal{S}_{\mathrm{R}}(\Sigma)$ was made by DeWitt (1970).

In the open case, one can perform a 'one-point compactification', $\bar{\Sigma} := \Sigma \cup \{\infty\}$. The corresponding superspace is then defined as

$$\mathcal{S}(\Sigma) := \text{Riem } \bar{\Sigma}/\mathcal{D}_{\mathrm{F}}(\bar{\Sigma}), \tag{4.93}$$

where the $\mathcal{D}_{\mathrm{F}}(\bar{\Sigma})$ are all diffeomorphisms that fix the frames at infinity. The open and the closed case are closer to each other than expected. One can show (Fischer 1986) that $\mathcal{S}_{\mathrm{R}}(\bar{\Sigma})$ and $\mathcal{S}(\Sigma)$ are diffeomorphic. For this reason, one can restrict topological investigations to the former space (Giulini 1995*a*).

The DeWitt metric G^{abcd} (see (4.25)), plays the role of a metric on Riem Σ,

$$G(l,k) := \int_{\Sigma} \mathrm{d}^3x \; G^{abcd} l_{ab} k_{cd}, \tag{4.94}$$

where l and k denote tangent vectors at $h \in$ Riem Σ. Due to its symmetry properties, it can formally be considered as a symmetric 6×6 matrix at each space point (DeWitt 1967*a*). At each point, this matrix can therefore be diagonalized, and the signature turns out to read

$$\mathrm{diag}(-,+,+,+,+,+).$$

A quantity with similar properties is known in elasticity theory—the fourth-rank elasticity tensor c^{abcd} possesses the same symmetry properties as DeWitt's metric and, therefore, has (in three spatial dimensions) 21 independent components; see e.g. Marsden and Hughes (1983).

It must be emphasized that the negative sign in DeWitt's metric has nothing to do with the Lorentzian signature of space–time; in the Euclidean case, the minus sign stays and only the relative sign between the potential and kinetic terms will change ($\sigma = 1$ instead of -1 in (4.24)). Due to the presence of this minus sign, the kinetic term for the gravitational field is *indefinite*. To gain a deeper understanding of its meaning, consider the following class of (generalized) DeWitt metrics which exhaust (up to trivial transformations) the set of all ultralocal metrics (DeWitt 1967*a*; cf. also Schmidt 1990):

$$G^{abcd}_{\beta} = \frac{\sqrt{h}}{2}\left(h^{ac}h^{bd} + h^{ad}h^{bc} - 2\beta h^{ab}h^{cd}\right), \tag{4.95}$$

where β is any real number (the GR value is $\beta = 1$). Its inverse is then given by

$$G^{\alpha}_{abcd} = \frac{1}{2\sqrt{h}}\left(h_{ac}h_{bd} + h_{ad}h_{bc} - 2\alpha h_{ab}h_{cd}\right), \tag{4.96}$$

where

$$\alpha + \beta = 3\alpha\beta \tag{4.97}$$

(in GR, $\alpha = 0.5$).[22] What would be the meaning of the constraints $\mathcal{H}^{\mathrm{g}}_{\perp}$ and $\mathcal{H}^{\mathrm{g}}_{a}$ if the generalized metric (4.95) were used? Section 4.1 shows that for $\beta \neq 1$ the principle of path independence must be violated, since the GR value $\beta = 1$ follows uniquely from this principle. The theories defined by these more generalized constraints thus cannot correspond to foliation-invariant ('covariant') theories at the Lagrangian level. In a sense, these would be genuinely Hamiltonian theories. In fact, the model of 'Hořava–Lifshits gravity' mentioned in Section 2.2.2 uses a DeWitt metric that corresponds to a non-covariant theory (Hořava 2009).

One can perform the following coordinate transformation in Riem Σ:

$$\tau = 4\sqrt{|\beta - 1/3\,|}\;h^{1/4}\;, \quad \tilde{h}_{ab} = h^{-1/3}h_{ab}, \tag{4.98}$$

thus decomposing the three-metric into a 'scale part' τ and a 'conformal part' $\tilde{h}_{ab}$. The 'line element' in Riem Σ can then be written as

$$G^{abcd}_{\beta}\mathrm{d}h_{ab}\otimes\mathrm{d}h_{cd} = -\mathrm{sgn}\left(\beta - 1/3\right)\mathrm{d}\tau\otimes\mathrm{d}\tau + \frac{\tau^2}{16|\beta - 1/3|}\mathrm{tr}\left(\tilde{h}^{-1}\mathrm{d}\tilde{h}\otimes\tilde{h}^{-1}\mathrm{d}\tilde{h}\right) \tag{4.99}$$

(for the inverse metric in (4.69), one must use α as in (4.96).) It is evident that the line element becomes degenerate for $\beta = 1/3$ (corresponding to $\alpha \to \infty$); for $\beta < 1/3$ it becomes positive definite, whereas for $\beta > 1/3$ it is indefinite (this includes the GR case).

At each space point, G^{abcd}_{β} can be considered as a metric in the space of symmetric positive definite 3×3 matrices, which is isomorphic to $\mathbb{R}^6$. Thus,

$$\mathbb{R}^6 \cong \mathrm{GL}(3,\mathbb{R})/\mathrm{SO}(3) \cong \mathrm{SL}(3,\mathbb{R})/\mathrm{SO}(3) \times \mathbb{R}^{+}. \tag{4.100}$$

The $\tilde{h}_{ab}$ are coordinates on $\mathrm{SL}(3,\mathbb{R})/\mathrm{SO}(3)$, and τ is the coordinate on $\mathbb{R}^{+}$. The relation (4.100) corresponds to the form (4.99) of the line element. All structures on $\mathrm{SL}(3,\mathbb{R})/\mathrm{SO}(3) \times \mathbb{R}^{+}$ can be transferred to Riem Σ, since G^{abcd}_{β} is ultralocal.

One can give an interpretation of one consequence of the signature change that occurs for $\beta = 1/3$ in (4.99). For this, one calculates the acceleration of the three-volume $V = \int \mathrm{d}^3x\,\sqrt{h}$ (assuming it is finite) for $N = 1$. After some calculation, one finds the expression (Giulini and Kiefer 1994)

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\int \mathrm{d}^3x\,\sqrt{h} = -3(3\alpha - 1)\int \mathrm{d}^3x\,\sqrt{h}\left(\frac{2}{3}\,{}^{(3)}\!R - 2\Lambda - 16\pi G\left[\mathcal{H}_{\mathrm{m}} - \frac{1}{3}h^{ab}\frac{\partial\mathcal{H}_{\mathrm{m}}}{\partial h^{ab}}\right]\right). \tag{4.101}$$

[22] In d space dimensions, one has $\alpha + \beta = d\alpha\beta$.

We call gravity 'attractive' if the sign in front of the integral on the right-hand side is negative. This is because then:

1. A positive ${}^{(3)}R$ contributes with a negative sign and leads to a *de*celeration of the three-volume.
2. A positive cosmological constant acts repulsively.
3. In the coupling to matter, an overall sign change corresponds to a sign change in G.[23]

In the Hamiltonian constraint (4.69), the inverse metric (4.96) enters. The critical value separating the positive definite from the indefinite case is thus $\alpha = 1/3$. One therefore recognizes that there is an intimate relation between the signature of the DeWitt metric and the attractivity of gravity: only for an indefinite signature is gravity attractive. From observations (primordial helium abundance), one can estimate (Giulini and Kiefer 1994) that

$$0.4 \lesssim \alpha \lesssim 0.55. \tag{4.102}$$

This is, of course, in accordance with the GR value $\alpha = 0.5$.

We return now to the case of GR. The discussion so far has concerned the metric on Riem Σ, given by the DeWitt metric (4.25). Does this metric also induce a metric on superspace (Giulini 1995*b*)? In Riem Σ, one can distinguish between 'vertical' and 'horizontal' directions. The vertical directions are the directions along the orbits generated by the three-dimensional diffeomorphisms. Metrics along a given orbit describe the same geometry. Horizontal directions are defined as being orthogonal to the orbits, where orthogonality holds with respect to the DeWitt metric. Since the latter is indefinite, the horizontal directions may also contain vertical directions (this happens in the 'light-like' case for zero norm). Calling V_h and H_h the vertical and horizontal subspaces, respectively, with respect to a given metric h_{ab}, one can show that:

1. If $V_h \cap H_h = \{0\}$, then G^{abcd} can be projected to the horizontal subspace where it defines a metric.
2. If $V_h \cap H_h \neq \{0\}$, then there exist critical points in $\mathcal{S}(\Sigma)$ where the projected metric changes signature.

The situation is illustrated in Fig. 4.3.

The task is then to classify the regions in Riem Σ according to these two cases. There exist some partial results (Giulini 1995*b*). For metrics obeying ${}^{(3)}R_{ab} = \lambda h_{ab}$ ('Einstein metrics'), one has $V_h \cap H_h = \{0\}$ and, consequently, a metric on $\mathcal{S}(\Sigma)$ exists. Moreover, one can show that for ${}^{(3)}R > 0$ (topologically, the case of a three-sphere) there remains only *one negative direction* in H_h out of the infinitely many negative directions in the DeWitt metric (in V_h, there are infinitely many negative directions). The kinetic term in the Hamiltonian constraint is then of a truly hyperbolic nature (in contrast to the general hyperbolicity in the pointwise sense). This holds in particular in the vicinity of closed Friedmann universes, which are therefore distinguished in this respect. A perturbation around homogeneity and isotropy has exhibited this property explicitly (Halliwell and Hawking 1985). It is an open question 'how far' one has to

[23]This does not, of course, mean that there is a change of sign in all terms where G appears.

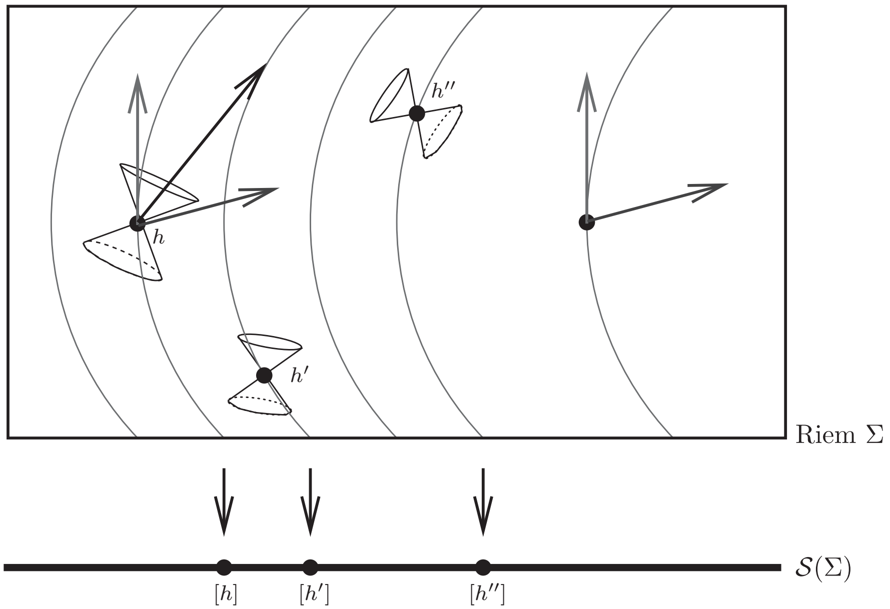

Fig. 4.3 The space Riem Σ, fibred by the orbits of Diff Σ (curved vertical lines). Tangent directions to these orbits are called 'vertical' (arrow pointing upwards), and the G-orthogonal directions are called 'horizontal' (arrow pointing to the right). Horizontal and vertical directions intersect whenever the 'hyper-lightcone' touches the vertical directions, as at point h'. At h, h', and h'' the vertical direction is depicted as time-like, light-like, and space-like, respectively. Hence $[h']$ corresponds to a transition point where the signature of the metric in superspace changes. Reprinted from Giulini and Kiefer (2007) with kind permission from Springer Science+Business Media B.V.

go from the hyperbolic case near ${}^{(3)}R > 0$ in order to reach the first points where the signature changes.

For Ricci-negative metrics[24] (i.e. all eigenvalues of ${}^{(3)}R_{ab}$ are negative), one finds that $V_h \cap H_h = \{0\}$, and that the projected metric on superspace contains infinitely many plus and minus signs. For flat metrics, one has $V_h \cap H_h \neq \{0\}$. Some of these results can be explicitly confirmed in Regge calculus; cf. Section 2.2.6. For more information about superspace, the reader may wish to consult Giulini (2009) and the references therein.

4.3 Canonical gravity with connections and loops

4.3.1 The canonical variables

One of the key ingredients in the canonical formalism is the choice of the symplectic structure, that is, the choice of the canonical variables. In the previous sections, we have chosen the three-metric h_{ab} and its momentum p^{cd}. In this section, we shall

[24]Any Σ admits such metrics.

introduce different variables introduced by Ashtekar (1986) following earlier work by Sen (1982). These 'new variables' will exhibit their main power in the quantum theory; see Chapter 6. Since they are analogous to Yang–Mills variables (using connections), the name 'connection dynamics' is also used. A more detailed introduction to these variables can be found in Ashtekar (1988, 1991) and—taking more recent developments into account—Thiemann (2007). An elementary introduction can be found in Gambini and Pullin (2011).

In their original version, Sen (1982) and Ashtekar (1986) made use of spinors in three-dimensional space. This was partly motivated by a desire to establish a connection with twistors (Penrose 1975), where complex numbers play a fundamental role. Later it was recognized that a spinorial representation is not needed and that it is, in fact, more natural to work with vector fields. This is also the path that we shall follow.

The first step in the construction of the 'new variables' consists in the introduction of *triads* (or dreibeine). These will play the role of the canonical momentum. Like the tetrads (vierbeine) used in Section 1.1.4, they are given by variables $e_i^a(x)$ that define an orthonormal basis at each space point. Here, $a = 1, 2, 3$ is the usual space index (referring to the tangent space $T_x(\Sigma)$ at x) and $i = 1, 2, 3$ are internal indices labelling the vectors. The position of the internal indices is arbitrary. One has the orthonormality condition

$$h_{ab} e_i^a e_j^b = \delta_{ij}, \tag{4.103}$$

from which one gets

$$h^{ab} = \delta^{ij} e_i^a e_j^b \equiv e_i^a e_i^b. \tag{4.104}$$

This introduces an SO(3) (or SU(2)) symmetry into the formalism, since the metric is invariant under local rotations of the triad. Associated with $e_i^a(x)$ is an orthonormal frame in the cotangent space $T_x^*(\Sigma)$, denoted by $e_a^i(x)$ (a basis of one-forms). It obeys

$$e_a^i e_j^a = \delta_j^i \ , \quad e_a^i e_i^b = \delta_a^b. \tag{4.105}$$

The three-dimensional formalism using triads can be obtained from the corresponding space–time formalism by using the 'time gauge' $e_a^0 = -n_a = N t_{,a}$ (Schwinger 1963) for the one-forms.

The variable of interest is not the triad itself, but its densitized version (because it will play the role of the momentum),

$$E_i^a(x) := \sqrt{h}(x) e_i^a(x), \tag{4.106}$$

where, from (4.103), one has $\sqrt{h} = |\det(e_a^i)|$. We note that one has the freedom to allow for different orientations of the triad (right-handed or left-handed). One recognizes that, in view of (4.104), the triad (4.106) is related to the *inverse* three-metric h^{ab}.

To find the canonically conjugate quantity, consider first the extrinsic curvature in the form

$$K_a^i(x) := K_{ab}(x) e^{bi}(x), \tag{4.107}$$

where $K_{ab}(x)$ denotes the previous expression for the extrinsic curvature; cf. (4.45) and (4.48). One can show that K_a^i is canonically conjugate to E_i^a,

$$K_a^i \delta E^{ia} = \frac{K_{ab}}{2\sqrt{h}} \delta \left(E^{ia} E^{ib} \right) = \frac{K_{ab}}{2\sqrt{h}} \left(h \delta h^{ab} + h^{ab} \delta h \right)$$
$$= -\frac{\sqrt{h}}{2} \left(K^{ab} - K h^{ab} \right) \delta h_{ab} = -8\pi G p^{ab} \delta h_{ab},$$

where (4.104) has been used and $\delta h = -h h_{cd} \delta h^{cd}$. The SO(3) rotation connected with the introduction of the triads is generated by the constraints

$$\mathcal{G}_i(x) := \epsilon_{ijk} K_a^j(x) E^{ka}(x) \approx 0, \tag{4.108}$$

which have the structure of '$\mathbf{x} \times \mathbf{p}$' (the generator of rotations). They are called 'Gauss constraints'. Their presence also guarantees the symmetry of K_{ab}. (This can be seen by inserting (4.107) into (4.108), multiplying by ϵ^{ilm}, and contracting.)

An arbitrary vector field can be decomposed with respect to the triad as

$$v^a = v^i e_i^a. \tag{4.109}$$

The covariant derivative with respect to internal indices is defined by

$$D_a v^i = \partial_a v^i + \omega_a{}^i{}_j v^j, \tag{4.110}$$

where the $\omega_a{}^i{}_j$ are the components of the spatial 'spin connection'; cf. also Section 1.1.4, where the space–time spin connection $\omega_{\mu mk}$ was used. One has the following relation between the spin connection and the Levi–Civita connection:

$$\omega_a{}^i{}_j = \Gamma^i_{kj} e^k_a, \tag{4.111}$$

where Γ^i_{kj} are the components of the Levi–Civita connection with respect to the triads. The usual coordinate components are found from

$$\Gamma^i_{kj} = e^d_k e^f_j e^i_c \Gamma^c_{df} - e^d_k e^f_j \partial_d e^i_f. \tag{4.112}$$

Inserting (4.111) and (4.112) into

$$D_a e^i_b = \partial_a e^i_b - \Gamma^c_{ab} e^i_c + \omega_a{}^i{}_j e^j_b,$$

one finds the covariant constancy of the triads,

$$D_a e^i_b = 0, \tag{4.113}$$

in analogy to $D_a h_{bc} = 0$. Parallel transport is defined by

$$\mathrm{d}v^i = -\omega_a{}^i{}_j v^j \mathrm{d}x^a.$$

Defining[25]

$$\Gamma^i_a = -\tfrac{1}{2}\, \omega_{ajk} \epsilon^{ijk}, \tag{4.114}$$

[25] ϵ_{ijk} is always the invariant tensor density here; that is, $\epsilon_{123} = 1$, etc.

this parallel transport corresponds to the infinitesimal rotation of the vector v^i by an angle

$$\delta\omega^i = \Gamma^i_a \, \mathrm{d}x^a, \tag{4.115}$$

that is,

$$\mathrm{d}v^i = \epsilon^i_{jk} v^j \delta\omega^k.$$

(Recall that for an orthonormal frame we have $\omega_{ajk} = -\omega_{akj}$.) From (4.113), one finds

$$\partial_{[a} e^i_{b]} = -\omega_{[a\ j}^{\ i} e^j_{b]} = -\epsilon^i_{jk} \Gamma^j_{[a} e^k_{b]}. \tag{4.116}$$

Parallel transport around a closed loop yields

$$\mathrm{d}v^i = -R^i_{jab} v^j \, \mathrm{d}x^a \, \mathrm{d}x^b \equiv \epsilon^i_{jk} v^j \delta\omega^k,$$

where R^i_{jab} are the components of the curvature two-form. The angle $\delta\omega^k$ can be written as

$$\delta\omega^k = -R^k_{ab} \, \mathrm{d}x^a \, \mathrm{d}x^b, \tag{4.117}$$

with $R^k_{ab} \epsilon^i_{jk} := R^i_{jab}$. The curvature components R^i_{ab} obey (from Cartan's second equation)

$$R^i_{ab} = 2\partial_{[a} \Gamma^i_{b]} + \epsilon^i_{jk} \Gamma^j_a \Gamma^k_b \tag{4.118}$$

and the 'cyclic identity'

$$R^i_{ab} e^b_i = 0. \tag{4.119}$$

The curvature scalar is given by

$$R[e] = -R^i_{ab} \epsilon_i^{\ jk} e^a_j e^b_k = -R^j_{kab} e^a_j e^{bk}. \tag{4.120}$$

The triad e^a_i (and similarly K^i_a) contains nine variables instead of the six variables of h_{ab}. The Gauss constraints (4.108) reduce the number from nine to six again.

The formalism presented up to now was known long ago. The progress achieved by Ashtekar (1986) consists in the second step—the mixing of E^a_i and K^i_a into a connection A^i_a. This is defined by

$$GA^i_a(x) = \Gamma^i_a(x) + \beta K^i_a(x), \tag{4.121}$$

where the 'Barbero–Immirzi' parameter β (Barbero 1995, Immirzi 1997) can at this stage be any (non-vanishing) complex number. It must be emphasized that the product GA^i_a has the dimension of an inverse length (like a Yang–Mills connection), but A^i_a itself has dimension mass over length squared. Therefore, GA is the relevant quantity ($\int GA \, \mathrm{d}x$ is dimensionless). The important fact is that A^i_a and $E^b_j/8\pi\beta$ are canonically conjugate variables,

$$\{A^i_a(x), E^b_j(y)\} = 8\pi\beta \delta^i_j \delta^b_a \delta(x, y). \tag{4.122}$$

In addition, one has

$$\{A^i_a(x), A^j_b(y)\} = 0. \tag{4.123}$$

In the following, A^i_a will be considered as the new *configuration* variable and E^b_j will be the corresponding canonical momentum.

4.3.2 Discussion of the constraints

The task now is to rewrite all constraints in terms of the new variables. We start with the Gauss constraints, which are only present due to the use of triads instead of metrics and the associated SO(3) redundancy. Using (4.121), one finds after some straightforward calculations[26]

$$\mathcal{G}_i = \partial_a E^a_i + G\epsilon_{ijk} A^j_a E^{ka} =: \mathcal{D}_a E^a_i \approx 0. \tag{4.124}$$

We are evidently justified in calling the constraints 'Gauss constraints', because they have a form similar to the Gauss constraints of Yang–Mills theories; cf. (4.33). In (4.124), we have also defined the covariant derivative $\mathcal{D}_a$ associated with A^i_a. Its associated curvature is

$$F^i_{ab} = 2G\partial_{[a} A^i_{b]} + G^2 \epsilon_{ijk} A^j_a A^k_b. \tag{4.125}$$

The Gauss constraints generate transformations similar to those in the Yang–Mills case:

$$\delta E^a_j(x) = \int \mathrm{d}y\ \{E^a_j(x), \mathcal{G}_i(y)\}\xi^i(y) = -8\pi\beta G\epsilon_{ijk} E^{ka}\xi^i$$

and

$$\delta A^i_a(x) = \int \mathrm{d}y\ \{A^i_a(x), \mathcal{G}_i(y)\}\xi^i(y) = -8\pi\beta\mathcal{D}_a\xi^i.$$

Sometimes it is also convenient to introduce su(2)-valued matrices

$$E^a = \tau_i E^a_i\ , \quad A_a = \tau_i A^i_a, \tag{4.126}$$

where $\tau_i = \mathrm{i}\sigma_i/2$, with σ_i being the Pauli matrices. Under an SU(2) transformation g, one then has

$$E^a \to gE^a g^{-1}\ , \quad A_a \to g(A_a + \partial_a)g^{-1}. \tag{4.127}$$

The next step is to rewrite the original gravitational constraints (4.69) and (4.70) in terms of the new variables. Introducing $\tilde{\mathcal{H}}_\perp = -8\pi G\beta^2\mathcal{H}^{\mathrm{g}}_\perp$ (plus terms proportional to the Gauss constraints) and $\tilde{\mathcal{H}}_a = -8\pi G\beta\mathcal{H}^{\mathrm{g}}_a$ (plus terms proportional to the Gauss constraints), the new form of the constraints is

$$\begin{aligned}\tilde{\mathcal{H}}_\perp = &-\frac{\sigma}{2}\frac{\epsilon^{ijk}F_{abk}}{\sqrt{|\det E^a_i|}}E^a_i E^b_j \\ &+\frac{\beta^2\sigma - 1}{\beta^2\sqrt{|\det E^a_i|}}E^a_{[i}E^b_{j]}(GA^i_a - \Gamma^i_a)(GA^j_b - \Gamma^j_b) \approx 0\end{aligned} \tag{4.128}$$

and

$$\tilde{\mathcal{H}}_a = F^i_{ab}E^b_i \approx 0. \tag{4.129}$$

Equation (4.129) has the form of the cyclic identity (4.119) with R^i_{ab} replaced by F^i_{ab}. If applied on A^i_a, the constraint (4.129) yields a transformation that can be written as a sum of a gauge transformation and a pure diffeomorphism.

[26]The constraint is also redefined through multiplication by β.

As in Section 4.1, one has $\sigma = -1$ for the Lorentzian and $\sigma = 1$ for the Euclidean case. One recognizes that (4.128) can be considerably simplified by choosing $\beta = \mathrm{i}$ (or $\beta = -\mathrm{i}$) for the Lorentzian and $\beta = 1$ (or $\beta = -1$) for the Euclidean case, because then the second term vanishes. The potential term has disappeared, leading to a situation resembling the strong-coupling limit discussed at the end of Section 4.2.3 (see also Section III.4 of Ashtekar (1988)). In fact, the original choice was $\beta = \mathrm{i}$ for the relevant Lorentzian case. Then,

$$2\sqrt{|\det E_i^a|}\,\tilde{\mathcal{H}}_\perp = \epsilon^{ijk} F_{abk} E_i^a E_j^b \approx 0.$$

This leads to a *complex* connection A_a^i (see (4.121)) and makes it necessary to implement reality conditions in order to recover GR—a task that seems impossible to achieve in the quantum theory. However, the choice $\beta = \pm\mathrm{i}$ is geometrically preferred (Rovelli 1991*a*, Penrose 2005); A_a^i is then the three-dimensional projection of a four-dimensional self-dual spin connection A_μ^{IJ},

$$A_\mu^{IJ} = \omega_\mu^{IJ} - \tfrac{1}{2}\,\mathrm{i}\epsilon^{IJ}{}_{MN}\omega_\mu^{MN}. \qquad (4.130)$$

Self-duality means that A_μ^{IJ} obeys the relation

$$A_\mu^{IJ} = -\frac{\mathrm{i}}{2}\epsilon^{IJ}_{MN} A_\mu^{MN}.$$

It turns out that the curvature $F_{\mu\nu}^{IJ}$ of the self-dual connection is the self-dual part of the Riemann curvature.

Values of the Barbero–Immirzi parameter other than $\beta = \pm\mathrm{i}$ do not seem to be of geometrical significance.[27] To avoid the problems with the reality conditions, however, preference has been given to real values. Barbero (1995) has chosen $\beta = -1$ for the Lorentzian case (which does not seem to have a special geometrical significance), so the Hamiltonian constraint reads[28]

$$\tilde{\mathcal{H}}_\perp = \frac{\epsilon^{ijk} E_i^a E_j^b}{2\sqrt{h}}(F_{abk} - 2R_{abk}) \approx 0. \qquad (4.131)$$

An alternative form using (4.126) is

$$\tilde{\mathcal{H}}_\perp = \frac{1}{\sqrt{h}}\mathrm{tr}\left((F_{ab} - 2R_{ab})[E^a, E^b]\right). \qquad (4.132)$$

The constraint algebra (3.84)–(3.86) remains practically unchanged, but one should keep in mind that the constraints have been modified by a term proportional to the Gauss constraints $\mathcal{G}_i$; cf. also (4.28). In addition, one has, of course, the relation for the generators of SO(3),

$$\{\mathcal{G}_i(x), \mathcal{G}_j(y)\} = \epsilon_{ij}{}^k \mathcal{G}_k(x)\delta(x, y). \qquad (4.133)$$

Following Thiemann (1996), we shall now rewrite the Hamiltonian constraint in a way that will turn out to be very useful in the quantum theory (Section 6.3). This

[27] The Barbero–Immirzi parameter arises naturally if torsion is present; cf. Baekler and Hehl (2011).

[28] In fact, the physics of black holes seems to prefer a very peculiar value for β; see (7.72) in Chapter 7.

will be achieved by expressing the Hamiltonian through Poisson brackets involving geometric quantities (area and volume). Consider for this purpose first the 'Euclidean part'[29] of $\tilde{\mathcal{H}}_\perp$,

$$\mathcal{H}_{\mathrm{E}} = \frac{\mathrm{tr}(F_{ab}[E^a, E^b])}{\sqrt{h}}. \tag{4.134}$$

(As we have discussed above, for $\beta = \mathrm{i}$ only this term remains.) Recalling (4.126), one finds

$$[E^a, E^b]_i = -\sqrt{h}\epsilon^{abc} e_c^i. \tag{4.135}$$

Here, use of the 'determinant formula'

$$(\det e_i^d)\epsilon^{abc} = e_i^a e_j^b e_k^c \epsilon^{ijk}$$

has been made. From the expression for the volume,

$$V = \int_\Sigma \mathrm{d}^3x\ \sqrt{h} = \int_\Sigma \mathrm{d}^3x\ \sqrt{|\det E_i^a|}, \tag{4.136}$$

one gets $2\delta V/\delta E_i^c(x) = e_c^i(x)$ and therefore

$$\frac{[E^a, E^b]_i}{\sqrt{h}} = -2\epsilon^{abc}\frac{\delta V}{\delta E_i^c} = -2\frac{\epsilon^{abc}}{8\pi\beta}\{A_c^i, V\}. \tag{4.137}$$

This yields for $\mathcal{H}_{\mathrm{E}}$ the expression

$$\mathcal{H}_{\mathrm{E}} = -\frac{1}{4\pi\beta}\epsilon^{abc}\,\mathrm{tr}(F_{ab}\{A_c, V\}). \tag{4.138}$$

Thiemann (1996) also considered the integrated trace of the extrinsic curvature,

$$T := \int_\Sigma \mathrm{d}^3x\ \sqrt{h}K = \int_\Sigma \mathrm{d}^3x\ K_a^i E_i^a,$$

for which one gets

$$\{A_a^i(x), T\} = 8\pi\beta K_a^i(x). \tag{4.139}$$

For $H_{\mathrm{E}} := \int \mathrm{d}^3x\ \mathcal{H}_{\mathrm{E}}$, one finds, using (4.116),

$$\{H_{\mathrm{E}}, V\} = 8\pi\beta^2 GT. \tag{4.140}$$

One now considers the following sum (written here for general β),

$$\tilde{\mathcal{H}}_\perp + \frac{1-\beta^2(\sigma+1)}{\beta^2}\mathcal{H}_{\mathrm{E}} = \frac{\beta^2\sigma - 1}{2\beta^2|\det E_i^a|}\left(F_{ab}^i - R_{ab}^i\right)[E^a, E^b]_i. \tag{4.141}$$

The reason for using this combination is to get rid of the curvature term. From (4.118) and using (4.121), one can write

[29] The name stems from the fact that, for $\beta = 1$ and the Euclidean signature $\sigma = 1$, this is already the full Hamiltonian $\tilde{\mathcal{H}}_\perp$.

$$R^i_{ab} = F^i_{ab} + \beta^2 \epsilon^i_{jk} K^j_a K^k_b + 2\beta \mathcal{D}_{[b} K^i_{a]}.$$

With the help of (4.139) and (4.135), one then finds, after some straightforward manipulations,

$$\tilde{\mathcal{H}}_\perp = -\frac{1-\beta^2(\sigma+1)}{\beta^2}\mathcal{H}_{\mathrm{E}} + \frac{\beta^2\sigma - 1}{2(4\pi\beta)^3}\epsilon^{abc}\ \mathrm{tr}\left(\{A_a, T\}\{A_b, T\}\{A_c, V\}\right). \tag{4.142}$$

This will serve as the starting point for the discussion of the quantum Hamiltonian constraint in Section 6.3. The advantage of this formulation is that $\tilde{\mathcal{H}}_\perp$ is fully expressed through Poisson brackets with geometric operators.

4.3.3 Loop variables

An alternative formulation that is closely related to the variables discussed in the last subsections employs so-called ‘loop variables’, introduced by Rovelli and Smolin (1990). This is presently the most frequently used formulation in the quantum theory (Chapter 6). Consider for this purpose a closed loop on Σ, that is, a continuous piecewise analytic map from the interval $[0,1]$ to Σ,

$$\alpha : [0,1] \to \Sigma\ , \quad s \mapsto \{\alpha^a(s)\}. \tag{4.143}$$

The *holonomy* $U[A,\alpha]$ corresponding to $A_a = A^i_a \tau_i$ along the curve α is given by

$$\begin{aligned} &U[A,\alpha](s) \in \mathrm{SU}(2)\ , \quad U[A,\alpha](0) = \mathbb{I}, \\ &\frac{\mathrm{d}}{\mathrm{d}s}U[A,\alpha](s) - G A_a(\alpha(s))\dot{\alpha}^a(s) U[A,\alpha](s) = 0, \end{aligned} \tag{4.144}$$

where $\dot{\alpha}^a(s) := \mathrm{d}\alpha^a/\mathrm{d}s$ (the tangential vector to the curve) and $U[A,\alpha](s)$ is a shorthand for $U[A,\alpha](0,s)$. The formal solution for the holonomy reads

$$U[A,\alpha](0,s) = \mathcal{P}\exp\left(G\int_\alpha A\right) \equiv \mathcal{P}\exp\left(G\int_0^s \mathrm{d}\tilde{s}\ \dot{\alpha}^a A^i_a(\alpha(\tilde{s}))\tau_i\right). \tag{4.145}$$

Here, $\mathcal{P}$ denotes path ordering, which is necessary because the A are matrices (as in Yang–Mills theories). One has, for example, for $s=1$,

$$\begin{aligned} &\mathcal{P}\exp\left(G\int_0^1 \mathrm{d}s\ A(\alpha(s))\right) \equiv U[A,\alpha] \\ &= \mathbb{I} + G\int_0^1 \mathrm{d}s\ A(\alpha(s)) + G^2\int_0^1 \mathrm{d}s \int_0^s \mathrm{d}t\ A(\alpha(t))A(\alpha(s)) + \ldots. \end{aligned}$$

We note that the A^i_a can be reconstructed uniquely if all holonomies are known (Giles 1981).

The holonomy is not yet gauge invariant with respect to SU(2) transformations. Under $g \in \mathrm{SU}(2)$, it transforms as

$$U[A,\alpha] \to U^g[A,\alpha] = gU[A,\alpha]g^{-1}.$$

Gauge invariance is achieved after performing the trace, thus arriving at the 'Wilson loop' known, for example, from lattice gauge theories,

$$\mathcal{T}[\alpha] = \mathrm{tr}\, U[A,\alpha]. \tag{4.146}$$

One can also define

$$\mathcal{T}^a[\alpha](s) = \mathrm{tr}\left[U[A,\alpha](s,s)E^a(s)\right], \tag{4.147}$$

where E^a is inserted at the point s of the loop. Analogously, one can define higher 'loop observables',

$$\mathcal{T}^{a_1\ldots a_N}[\alpha](s_1,\ldots,s_n),$$

by inserting E^a at the corresponding points described by the s-values. These loop observables obey a closed Poisson algebra called the *loop algebra*. One has, for example,

$$\{\mathcal{T}[\alpha], \mathcal{T}^a[\beta](s)\} = \Delta^a[\alpha,\beta(s)]\left(\mathcal{T}[\alpha\#\beta] - \mathcal{T}[\alpha\#\beta^{-1}]\right), \tag{4.148}$$

where

$$\Delta^a[\alpha,x] = \int \mathrm{d}s\ \dot{\alpha}^a(s)\delta(\alpha(s),x), \tag{4.149}$$

and β^{-1} denotes the loop β with the reversed direction. The right-hand side of (4.149) is only non-vanishing if α and β have an intersection at a point P; $\alpha\#\beta$ is then defined as starting from P, going through the loop α, then through β, and ending at P.

Of particular interest is the quantity

$$E[\mathcal{S},f] := \int_{\mathcal{S}} \mathrm{d}\sigma_a\ E^a_i f^i, \tag{4.150}$$

where $\mathcal{S}$ denotes a two-dimensional surface in Σ, and $f = f^i\tau_i$. The variable (4.150) describes the flux of E^a_i through the two-dimensional surface; it is the variable conjugate to the holonomy $U[A,\alpha]$. Note that both the holonomy and the flux are of a distributional nature because they have support on one-dimensional and two-dimensional submanifolds, respectively. These variables will be used in the quantum theory; see Chapter 6.

An important feature of the loop variables is their non-locality. It is basically this fact that leads to their successful quantization (Chapter 6). Non-local variables have been suggested before. In electrodynamics, perhaps the first instance where the line integral of the vector potential was suggested as the basic variable is Weiss (1938). He emphasized that this integral has an absolute significance, in contrast to the vector potential itself. A path-dependent variable for the gravitational field was suggested by Mandelstam (1962). However, his variable is a four-dimensional one containing the Riemann tensor and is unrelated to the three-dimensional loop variables employed here.

Further reading: Barbour *et al.* (2002), Hojman *et al.* (1976), Rovelli (2004).

5 Quantum geometrodynamics

5.1 The programme of canonical quantization

Given a classical theory, one cannot derive a unique 'quantum theory' from it. The only possibility is to 'guess' such a theory and to test it by experiment. For this purpose, sets of 'quantization rules' have been devised, which have turned out to be successful in the construction of quantum theories; one example is quantum electrodynamics. Strictly speaking, the task is to construct a quantum theory from its classical limit.

In Chapter 4, we developed a Hamiltonian formulation of GR. This is the appropriate starting point for a canonical quantization, which requires the definition of a configuration variable and its conjugate momentum. A special feature of GR is the fact, as is the case in all fully reparametrization-invariant systems, that the dynamics is entirely generated by constraints: the total Hamiltonian either vanishes as a constraint (for the spatially compact case) or consists solely of surface terms (in the asymptotically flat case). The central difficulty is thus, both conceptually and technically, the correct treatment of the quantum constraints, that is, the quantum version of the constraints (4.69) and (4.70) or their versions in the connection or loop representation.

In Chapter 3, we presented a general procedure for the quantization of constrained systems. Following Dirac (1964), a classical constraint is turned into a restriction on physically allowed wave functionals (see Section 3.1),

$$\mathcal{G}_a \approx 0 \longrightarrow \hat{\mathcal{G}}_a \Psi = 0. \tag{5.1}$$

At this stage, such a transition is only a heuristic recipe that has to be made more precise. Following Ashtekar (1991) and Kuchař (1993), we shall divide the 'programme of canonical quantization' into six steps, which will be presented here briefly and then implemented (or attempted to be implemented) in the following sections.

The **first step** consists in the identification of configuration variables and their momenta. In the language of geometric quantization (Woodhouse 1992), it is the choice of polarization. Together with the unit operator, these variables are called the 'fundamental variables' V_i. The implementation of Dirac's procedure is the translation of Poisson brackets into commutators for the fundamental variables, that is,

$$V_3 = \{V_1, V_2\} \longrightarrow \hat{V}_3 = -\frac{\mathrm{i}}{\hbar}[\hat{V}_1, \hat{V}_2]. \tag{5.2}$$

In the geometrodynamical formulation of GR (see Sections 4.1 and 4.2), the fundamental variables are, apart from the unit operator, the three-metric $h_{ab}(\mathbf{x})$ and its momentum $p^{cd}(\mathbf{x})$ (or, in the approach of reduced quantization, a subset of them; see Section 5.2 below). In the connection formulation of Section 4.3, one has the connection $A_a^i(\mathbf{x})$ and the densitized triad $E_j^b(\mathbf{y})$, while in the loop-space formulation one takes the holonomy $U[A, \alpha]$ and the flux of E_j^b through a two-dimensional surface.

In this chapter, we restrict ourselves to the quantization of the geometrodynamical formulation, while quantum connection dynamics and quantum loop dynamics will be discussed in Chapter 6. Application of (5.2) to (4.64) would yield

$$[\hat{h}_{ab}(\mathbf{x}), \hat{p}^{cd}(\mathbf{y})] = \mathrm{i}\hbar\delta^c_{(a}\delta^d_{b)}\delta(\mathbf{x}, \mathbf{y}), \tag{5.3}$$

plus vanishing commutators between all the metric components, and between all momentum components. Since p^{cd} is linearly related to the extrinsic curvature, describing the embedding of the three-geometry into the fourth dimension, the presence of the commutator (5.3) and the ensuing 'uncertainty relation' between intrinsic and extrinsic geometry means that the classical space–time picture has completely dissolved in quantum gravity. This is fully analogous to the disappearance of particle trajectories as fundamental concepts in quantum mechanics and constitutes one of the central interpretational ingredients of quantum gravity. The fundamental variables form a vector space that is closed under Poisson brackets and complete in the sense that every dynamical variable can be expressed as a sum of products of fundamental variables.

Equation (5.3) does not implement the positivity requirement $\det h > 0$ of the classical theory. But this would only definitely be a problem if (the smeared version of) $\hat{p}^{ab}$ were self-adjoint and its exponentiation were therefore a unitary operator, which could 'shift' the metric to negative values. An approach that attempts to implement the positivity requirement consistently is the programme of affine quantum gravity (Klauder 2010).

The **second step** addresses the quantization of a general variable, F, of the fundamental variables. Does the rule (5.2) still apply? As Dirac (1958*a*, p. 87) writes,[1]

> The strong analogy between the quantum P.B. ... and the classical P.B. ... leads us to make the assumption that the quantum P.B.s, or at any rate the simpler ones of them, have the same values as the corresponding classical P.B.s. The simplest P.B.s are those involving the canonical coordinates and momenta themselves ...

In fact, from general theorems of quantum theory (going back to Groenewald and van Hove), one knows that it is impossible to respect the transformation rule (5.2) in the general case while assuming an irreducible representation of the commutation rules; cf. Giulini (2003). In Dirac's quote, this is anticipated by the statement 'or at any rate the simpler ones of them'. This failure is related to the problem of 'factor ordering'. Therefore, additional criteria must be invoked to find the 'correct' quantization, such as the demand for 'Dirac consistency', to be discussed in Section 5.3.

The **third step** concerns the construction of an appropriate representation space, $\mathcal{F}$, for the dynamical variables, on which they should act as operators. We shall usually employ the functional Schrödinger picture, in which the operators act on wave functionals defined in an appropriate functional space. In the present case, the application of the Schrödinger picture can at best be understood in a formal sense.

[1] 'P.B.' stands for 'Poisson bracket'.

The implementation of the fundamental commutation relations (5.3) would be achieved by writing

$$\hat{h}_{ab}(\mathbf{x})\Psi\left[h_{ab}(\mathbf{x})\right] = h_{ab}(\mathbf{x}) \cdot \Psi\left[h_{ab}(\mathbf{x})\right], \tag{5.4}$$

$$\hat{p}^{cd}(\mathbf{x})\Psi\left[h_{ab}(\mathbf{x})\right] = \frac{\hbar}{\mathrm{i}}\frac{\delta}{\delta h_{cd}(\mathbf{x})}\Psi\left[h_{ab}(\mathbf{x})\right]. \tag{5.5}$$

These relations do not define self-adjoint operators, since there is no Lebesgue measure on Riem Σ (which would have to be invariant under translations in function space). Thus, one would not expect the fundamental relations (5.3) to be necessarily in conflict with $\det h > 0$. Other examples of the use of the functional Schrödinger picture have already been presented in Chapter 3.

The representation space $\mathcal{F}$ is only an auxiliary space: before the constraints are implemented, it does not necessarily contain only physical states. Therefore, neither does it have to be a Hilbert space, nor do operators acting on $\mathcal{F}$ have to be self-adjoint. It might even be inconsistent to demand that the constraints be self-adjoint operators on an auxiliary Hilbert space $\mathcal{F}$.

The **fourth step** consists in the implementation of the constraints. According to (5.1), one would implement the classical constraints $\mathcal{H}_\perp \approx 0$ and $\mathcal{H}_a \approx 0$ as

$$\mathcal{H}_\perp \Psi = 0, \tag{5.6}$$

$$\mathcal{H}_a \Psi = 0. \tag{5.7}$$

These are infinitely many equations (one equation at each space point), which we denote collectively by $\mathcal{H}_\mu \Psi = 0$. Only solutions to these 'quantum constraints' can be regarded as candidates for physical states. The solution space will be called $\mathcal{F}_0$. How the constraints (5.6) and (5.7) are written in detail depends on one's approach to the 'problem of time'; see Section 5.2. It is to be expected that the solution space will still be too large; as in quantum mechanics, one may have to impose further conditions on the wave functions, such as normalizability. This requirement is needed in quantum mechanics because of the probability interpretation, but it is far from clear whether this interpretation can be maintained in quantum gravity; cf. Chapter 10. But even if this is not required in quantum gravity, some other conditions may be needed. Ideally, thus, the physical space $\mathcal{F}_{\mathrm{phys}}$ on which wave functionals act should be a genuine subspace, $\mathcal{F}_{\mathrm{phys}} \subset \mathcal{F}_0 \subset \mathcal{F}$.

The **fifth step** concerns the role of observables. We have already mentioned in Section 3.1 that 'observables' are characterized by having weakly vanishing Poisson brackets with the constraints, $\{\mathcal{O}, \mathcal{G}_a\} \approx 0$. They have, a priori, nothing to do with concrete measurement prescriptions. In quantum mechanics, observables are associated in a somewhat vague manner with self-adjoint operators (only this concept is mathematically precise). In practice, however, only a few operators correspond in fact to quantities that are 'measured'. Only the latter represent 'beables' in the sense of John Bell, supposedly describing 'reality'; see Bell (2004).[2] For an operator corresponding to a classical observable satisfying $\{\mathcal{O}, \mathcal{H}_\mu\} \approx 0$, one would expect that in the quantum theory the relation

[2]These are quantities that are subject to decoherence; see Chapter 10.

$$[\hat{\mathcal{O}}, \hat{\mathcal{H}}_\mu]\Psi = 0 \tag{5.8}$$

will hold. For operators $\hat{F}$ with $[\hat{F}, \hat{\mathcal{H}}_\mu]\Psi \neq 0$, one would have $\hat{\mathcal{H}}_\mu(\hat{F}\Psi) \neq 0$. This is sometimes interpreted as meaning that the 'measurement' of the quantity related to this operator leads to a state that is no longer annihilated by the constraints, so that the wave function is thrown out of the solution space. It seems that this would only be a problem for a 'collapse' interpretation of quantum gravity, an interpretation that is highly unlikely to hold in quantum gravity; see Chapter 10. One could, of course, argue that only operators that qualify as operators are allowed by the formalism. This would, however, exclude important operators such as the fundamental variables themselves.

Since the classical Hamiltonian and diffeomorphism constraints differ from each other in their interpretation (Chapter 4), the same should hold for their quantum versions (5.6) and (5.7). This is, in fact, the case and will be discussed later in this chapter. The potential distinction between 'observables' and 'perennials' (see (4.84) and (4.85)) thus applies also to the quantum case.

The **sixth** (and last) **step** concerns the role of the physical Hilbert space (as previously mentioned in step 3). Do the observables have to be represented in some Hilbert space? If yes, which one? It certainly cannot be the auxiliary space $\mathcal{F}$, but it is unclear whether it is $\mathcal{F}_0$ or only $\mathcal{F}_{\text{phys}} \subset \mathcal{F}_0$. Moreover, it may turn out that only a construction with rigged Hilbert spaces (Gel'fand triples) is possible; this is in fact the case in the loop representation (Section 6.1.2).

A general method to deal with the construction of a physical Hilbert space in the quantization of constrained systems is the group averaging procedure. There, quantum states are averaged over the action of the invariance group. It has been shown that this method works and yields a unique $\mathcal{F}_{\text{phys}}$, at least for finite-dimensional compact Lie groups; cf. Giulini and Marolf (1999). (For an extension to non-compact Lie groups, see Louko (2006).) The situation for GR, where the constraint algebra is not a Lie algebra at all, remains unclear.

To represent all observables by self-adjoint operators in Hilbert space would be contradictory; for if $\hat{F}$ and $\hat{G}$ are self-adjoint observables, then the product $\hat{F}\hat{G}$ will again be an observable, but will no longer be self-adjoint, since

$$\left(\hat{F}\hat{G}\right)^\dagger = \hat{G}^\dagger \hat{F}^\dagger = \hat{G}\hat{F} \overset{\text{in general}}{\neq} \hat{F}\hat{G}.$$

Since, moreover, the fundamental variables h_{ab} and p^{cd} are not observables, one might, at this stage, decide to forget about this notion. The constraints themselves can most likely not all be represented as self-adjoint operators (Komar 1979). The 'Hilbert-space problem' is intimately connected with the 'problem of time' in quantum gravity, to which we shall now turn.

5.2 The problem of time

The concepts of time in GR and in quantum theory differ drastically from each other. As already remarked in Section 1.1, time in quantum theory is an external parameter (an absolute element of the theory), whereas GR is background-independent and thus cannot depend on an external time. A consistent theory of quantum gravity should, therefore, exhibit a novel concept of time. The history of physics has shown that new theories often entail a new concept of space and time (Ehlers 1973). The same should happen again with quantum gravity.

The absolute nature of time in quantum mechanics is crucial for its interpretation. Matrix elements are usually evaluated at fixed t, and the scalar product is conserved in time ('unitarity'). Unitarity expresses the conservation of the total probability. 'Time' is part of the classical background that, according to the Copenhagen interpretation, is needed for the interpretation of measurements. As we remarked at the end of Section 3.1, the introduction of a time operator in quantum mechanics is problematic. The time parameter t appears explicitly in the Schrödinger equation (3.14). Note that it comes together with the imaginary unit i, a fact that finds an explanation in the semiclassical approximation to quantum geometrodynamics (Section 5.4). The occurrence of the imaginary unit in the Schrödinger equation was discussed in an interesting correspondence between Ehrenfest and Pauli; see Pauli (1985, p. 127). Pauli pointed out that the use of complex wave functions can be traced back to the probability interpretation:[3]

> I now turn to the initially asked question about the necessity of at least *two* scalars for the de Broglie–Schrödinger waves. I claim that this necessity and thus also the imaginary unit come into play *through the search for an expression for the probability density W which satisfies conditions (1) and (2) and which does not contain the temporal derivatives of ψ.*

Conditions (1) and (2) are the non-negativity of W and its normalization to one, respectively.

Since GR is background-independent, there is no absolute time. Space–time influences material clocks in order to allow them to show proper time. The clocks, in turn, react on the metric and change the geometry. In this sense, the metric itself *is* a clock (Zeh 2007). A quantization of the metric can thus be interpreted as a quantization of the concept of time. Since the nature of time in quantum gravity is not yet clear—the classical constraints do not contain any time parameter—one speaks of the 'problem of time'. One can distinguish basically three possible solutions of this problem, as reviewed, in particular, by Isham (1993) and Kuchař (1992):

1. Choice of a concept of time *before* quantization.
2. Identification of a concept of time *after* quantization.
3. *'Timeless'* options.

The first two possibilities will be discussed in the following, while the third option will be addressed in Sections 5.3 and 5.4.

5.2.1 Time before quantization

In Section 4.2.3, the arguments of Baierlein *et al.* (1962) were presented, to the extent that the three-dimensional geometry in GR contains information about time. Motivated by the parametrized theories discussed in Chapter 3, one can attempt to perform

[3] Nun komme ich zur anfangs gestellten Frage über die Notwendigkeit von mindestens *zwei* reellen Skalaren bei den de Broglie–Schrödinger-Wellen. Ich behaupte, diese Notwendigkeit und damit auch die imaginäre Einheit kommt hinein *beim Suchen nach einem Ausdruck für die Wahrscheinlichkeitsdichte W, der die Forderungen (1) und (2) befriedigt und der die zeitlichen Ableitungen der ψ nicht enthält.*

a canonical transformation, aiming at an isolation of time from the 'true degrees of freedom'. Starting from the 'ADM variables' $h_{ab}(\mathbf{x})$ and $p^{cd}(\mathbf{x})$, one would like to perform the step

$$\big(h_{ab}(\mathbf{x}), p^{cd}(\mathbf{x})\big) \longrightarrow \big(X^A(\mathbf{x}), P_B(\mathbf{x}); \phi^r(\mathbf{x}), p_s(\mathbf{x})\big)\,, \tag{5.9}$$

where the $8\times\infty^3$ variables X^A and P_B $(A, B = 0, 1, 2, 3)$ are the 'embedding variables' and their canonical momenta, while the $4\times\infty^3$ variables ϕ^r and p_s $(r, s = 1, 2)$ denote the 'true' degrees of freedom of the gravitational field; cf. also Section 7.4. As already remarked in Section 4.2, GR is not equivalent to a deparametrized theory. Therefore, (5.9) is certainly non-unique and not valid globally (see in this context Hájíček and Kijowski 2000). The next step is the elimination of $4\times\infty^3$ of the $8\times\infty^3$ embedding variables by casting the classical constraints $\mathcal{H}_\mu \approx 0$ into the form[4]

$$\mathcal{H}_A := P_A(\mathbf{x}) + h_A(\mathbf{x}; X^B, \phi^r, p_s] \approx 0. \tag{5.10}$$

This is referred to as 'solving the constraints at the classical level' or 'reduced quantization' and corresponds in the case of particle systems to (3.9). As in Section 3.1, the remaining $4\times\infty^3$ variables are eliminated by inserting (5.10) into the action

$$S = \int \mathrm{d}t \int_\Sigma \mathrm{d}^3x\; \Big(P_A \dot{X}^A + p_r\dot{\phi}^r - N\mathcal{H}_\perp - N^a\mathcal{H}_a\Big)\,, \tag{5.11}$$

where all fields are functions of $\mathbf{x}$ and t, and going to the constraint hypersurface ('deparametrization'), yielding

$$S = \int \mathrm{d}t \int_\Sigma \mathrm{d}^3x\; \Big(p_r\dot{\phi}^r - h_A(\mathbf{x}; X_t^B, \phi^r, p_s]\dot{X}_t^A(\mathbf{x})\Big)\,, \tag{5.12}$$

where $\dot{X}_t^A(\mathbf{x})$ is now a prescribed function of t and $\mathbf{x}$, which must not be varied. This action corresponds to the action (3.12), in which the prescribed function is $t(\tau)$. The action (5.12) describes an ordinary canonical system with a 'true', that is, unconstrained Hamiltonian given by

$$H_{\text{true}}(t) = \int_\Sigma \mathrm{d}^3x\; h_A(\mathbf{x}; X_t^B, \phi^r, p_s]\dot{X}_t^A(\mathbf{x}). \tag{5.13}$$

One can derive from H_{true} Hamilton's equations of motion for ϕ^r and p_s. The variables $X_t^A(\mathbf{x})$ can only be interpreted as describing embeddings in a space–time *after* these equations (together with the choice of the lapse and shift) have been solved.

The constraint (5.10) can be quantized in a straightforward manner by introducing wave functionals $\Psi[\phi^r(\mathbf{x})]$, with the result

$$\mathrm{i}\hbar\frac{\delta\Psi[\phi^r(\mathbf{x})]}{\delta X^A(\mathbf{x})} = h_A\left(\mathbf{x}; X^B, \hat{\phi}^r, \hat{p}^s\right]\Psi[\phi^r(\mathbf{x})], \tag{5.14}$$

in which the X^A have not been turned into an operator. In this respect, the quantization is of a hybrid nature: momenta occurring linearly in the constraints are formally turned into derivatives, although the corresponding configuration variables stay

[4]'$(\mathbf{x}$' etc. means dependence on $\mathbf{x}$ as a function, while '$p_s]$' etc. means dependence on $p_s(\mathbf{x})$ as a functional.

classical—like the t in the Schrödinger equation. Equation (5.14) has the form of a local Schrödinger equation. Such an equation is usually called a 'Tomonaga–Schwinger equation'; strictly speaking, it consists of infinitely many equations with respect to the local 'bubble time' $X^A(\mathbf{x})$. We shall say more about such equations in Section 5.4. The main advantages of this approach to quantization, if it were to succeed, would be:

1. One would have already isolated at the classical level a time variable (here, 'embedding variables') that is external to the quantum system described by $\hat{\phi}^r$ and $\hat{p}_s$. The formalism would thus be similar to ordinary quantum field theory.
2. Together with such a distinguished notion of time would come a natural Hilbert-space structure and its ensuing probability interpretation.
3. One would consider observables to be any function of the 'genuine' operators $\hat{\phi}^r$ and $\hat{p}_s$. As in the linearized approximation (Chapter 2), the gravitational field would have two degrees of freedom.

On the other hand, one faces many problems:

1. *Multiple-choice problem* The canonical transformation (5.9) is certainly non-unique and the question arises which choice should be made. One would expect that different choices of 'time' would lead to non-unitarily connected quantum theories.
2. *Problem of deparametrization* One cannot identify the constraint surface of GR with that of a parametrized theory (Torre 1993*a*); that is, (5.9) cannot hold globally. More precisely, Torre showed that the phase space of GR cannot be identified with the phase space of a parametrized theory, because the former has singularities, while the latter has not.
3. *h_A problem* The 'true' Hamiltonian (5.13) depends on 'time', that is, on the embedding variables X^A. This dependence is expected to be very complicated (leading to square roots of operators, etc.), prohibiting in general a rigorous definition.
4. *X^A problem* In the classical theory, the 'bubble time' X^A describes a hypersurface in space–time *only after* the classical equations have been solved. Since no classical equations and therefore no space–time are available in the quantum theory, (5.14) has no obvious space–time interpretation. In particular, an operational treatment of time is unknown.
5. *Space–time problem* Writing $X^A = (T, X^i)$, the 'time' T would have to be a space–time scalar in order to be interpretable as a real time. This would mean that, although it is constructed from the canonical data h_{ab} and p^{cd} on Σ, it must vanish weakly with the Hamiltonian constraint,

$$\left\{T(\mathbf{x}), \int_\Sigma \mathrm{d}^3y\ \mathcal{H}_\perp(\mathbf{y})N(\mathbf{y})\right\} \approx 0, \tag{5.15}$$

for all lapse functions $N(\mathbf{y})$ with $N(\mathbf{x}) = 0$. Otherwise, one would get from two hypersurfaces Σ and Σ' crossing at $\mathbf{x}$ two different values of T, depending on whether the canonical data of Σ or of Σ' were used. The variable T would in this case have no use as a time variable. This problem is related to the fact that the algebra of hypersurface transformations does not coincide with space–time diffeomorphisms. A possible solution to the space–time problem can be obtained by using matter variables, for example the 'reference fluid' used by Brown and Kuchař (1995). The 'space–time problem' anticipates a space–time picture, which, however, is absent in quantum gravity. The space–time problem therefore refers mainly to the classical theory. It must also be noted that the embedding variables never commute with all constraints, as can be seen from the form (5.10) of the constraints $\mathcal{H}_A$. They are therefore never space–time scalars in the strict sense.

6. *Anomalies* Quantum anomalies may spoil the consistency of this approach; cf. Section 5.3.
7. *Problem of construction* The actual transformation (5.9) has been performed only in very special cases, for example linearized gravity, cylindrical gravitational waves, black holes, and dust shells (Chapter 7), and homogeneous cosmological models (Chapter 8).

In the full theory, concrete proposals for the canonical transformation (5.9) are rare. A possibility that was developed to a certain extent and that is applicable for compact spaces makes use of 'York's time' or 'extrinsic time', which is defined by

$$T\left(\mathbf{x}; h_{ab}, p^{cd}\right] = \frac{2}{3\sqrt{h}} p^{cd} h_{cd}, \quad P_T = -\sqrt{h}\ ; \tag{5.16}$$

cf. Al'tshuler and Barvinsky (1996) and the references therein. The existence of the transformation (5.16) does not contradict the theorem proved in Torre (1993*a*), to which reference was made above in the context of the global-time problem: first, (5.16) refers to only one of the four constraints, and second, no statement is made about the extent to which (5.16) is possible in a global and regular way.

Since $p^{cd}h_{cd} = -\sqrt{h}K/8\pi G$ (cf. (4.63)), T is proportional to the trace of the extrinsic curvature K. It is canonically conjugate to P_T. Note that T does not obey (5.15) and is thus not a space–time scalar. It has been shown[5] that the Hamiltonian constraint can be written in the form (5.10), that is, written as $P_T + h_T \approx 0$, where h_T is known to exist, but not known in explicit form, that is, not known as an explicit function of T and the remaining variables. From (5.16) it is clear that the 'true' Hamiltonian contains the three-dimensional volume as its dynamical part, that is,

$$H_{\text{true}} = \int \mathrm{d}^3x\ \sqrt{h} + \int \mathrm{d}^3x\ N^a \mathcal{H}_a. \tag{5.17}$$

The main problem with this approach towards the issue of time in quantum gravity is perhaps its closeness to a classical space–time picture. From various equations, such as (5.14), one gets the illusion that a space–time exists even in the quantum theory, although this cannot be the case, see (5.3). One can, therefore, conclude that attempting to identify time before quantization does not solve the problem of time in the general case, although it might help in special cases (Section 7.4).

[5]This involves a detailed study of the 'Lichnerowicz equation', a non-linear (but quasi-linear) elliptical equation for P_T, which under appropriate conditions possesses a unique solution (cf. Choquet-Bruhat and York 1980).

5.2.2 Time after quantization

Using the commutation rules (5.3) and their formal implementation (5.4) and (5.5) directly, one arrives at wave functionals $\Psi[h_{ab}(\mathbf{x})]$ defined on Riem Σ, the space of all three-metrics. This is the central kinematical quantity. The 'dynamics' must be implemented through the quantization of the constraints (4.69) and (4.70)—this is all that remains in the quantum theory. One then gets the following equations for the wave functional:

$$\hat{\mathcal{H}}^{\mathrm{g}}_{\perp}\Psi := \left(-16\pi G\hbar^2 G_{abcd}\frac{\delta^2}{\delta h_{ab}\delta h_{cd}} - \frac{\sqrt{h}}{16\pi G}\left({}^{(3)}R - 2\Lambda\right)\right)\Psi = 0, \qquad (5.18)$$

$$\hat{\mathcal{H}}^{\mathrm{g}}_{a}\Psi := -2D_b h_{ac}\frac{\hbar}{\mathrm{i}}\frac{\delta\Psi}{\delta h_{bc}} = 0. \qquad (5.19)$$

Equation (5.18) is called the *Wheeler–DeWitt equation* in honour of the work by DeWitt (1967*a*) and Wheeler (1968).[6] In fact, this is again infinitely many equations. The constraints (5.19) are called the quantum diffeomorphism (or momentum) constraints. Occasionally, both (5.18) and (5.19) are referred to as the Wheeler–DeWitt equations. In the presence of non-gravitational fields, these equations are augmented by the corresponding terms.

In the first years of its discussion, (5.18) was called the 'Einstein–Schrödinger equation'. As Karel Kuchař remarks (Kuchař 1973, p. 29), [7]

> The quantum version (5.18) of the super-Hamiltonian constraint was called the Einstein–Schrödinger equation by John Wheeler, and the Wheeler–DeWitt equation by others. Because I cannot recall any other physicists who would more strenuously object to the idea of quantizing gravity than Einstein and Schrödinger, and because I am not John Wheeler, I shall use the second name.

There are many problems associated with the Wheeler–DeWitt equation (5.18). An obvious problem is the 'factor-ordering problem': the precise form of the kinetic term is open—there could be additional terms proportional to $\hbar$ containing at most first derivatives in the metric. Since second functional derivatives at the same space point usually lead to undefined expressions such as $\delta(0)$, a regularization (and perhaps renormalization) scheme has to be employed, cf. Isham (1976). Connected with this is the potential presence of anomalies. The general discussion of these problems is continued in Section 5.3. Here we shall address again the problem of time and the related Hilbert-space problem. Since (5.18) does not have the structure of a local Schrödinger equation (5.14), the choice of Hilbert space is not clear a priori. It is not even clear whether a Hilbert space is needed at the full level of quantum gravity (see the remarks at the end of this subsection).

The first option for an appropriate Hilbert space is related to the use of a *Schrödinger-type inner product*, that is, the standard quantum-mechanical inner product as generalized to quantum field theory,

[6] According to his own account, DeWitt had already formulated this equation in the early 1960s; see DeWitt (1999). Equations of the form $H\Psi = 0$ can be found for gravity in Dirac (1959) and Bergmann (1966), but their explicit form is not given.

[7] The equation number in this quote has been adapted to the present numbering.

$$\langle\Psi_1|\Psi_2\rangle = \int_{\mathrm{Riem}\ \Sigma} \mathcal{D}\mu[h]\ \Psi_1^*[h]\Psi_2[h], \tag{5.20}$$

where h here is a shorthand for h_{ab}. It is known that such a construction is at best formal, since the measure $\mathcal{D}\mu[h]$ cannot be rigorously defined; that is, there is no Lebesgue measure in the functional case. The elementary operators $\hat{h}_{ab}$ and $\hat{p}^{cd}$ are at best formally self-adjoint with respect to this inner product.

Besides the lack of mathematical rigour, this inner product has other problems. The integration runs over all metric components, including potential unphysical ones (the constraints have not yet been imposed at this stage). This could lead to divergences—similar to an integration over t in the quantum-mechanical case—which one would have to cure by changing the measure, as was done by the introduction of the Faddeev–Popov determinant into path integrals (Section 2.2); cf. Woodard (1993) in this context. This is related to the fact that the product (5.20) is defined on the full space $\mathcal{F}$, not the solution space $\mathcal{F}_0 \subset \mathcal{F}$. It might be possible to turn these problems into a virtue by imposing a 'boundary condition' that physical states solve the constraints *and* lead to a finite inner product (5.20). Such a proposal can at least be implemented within simple models (cf. Chapter 8), but one would face the danger that in the full theory no such solutions might exist at all. An open problem is also the implementation of a probability interpretation in this context (recall that this is the major motivation for using this inner product in quantum mechanics). What does the probability to find a certain three-metric mean? The answer is unclear. It is possible that the Schrödinger inner product only makes sense in the semiclassical approximation (Section 5.4).

We have seen in Section 4.2 that the kinetic term of the Hamiltonian constraint is *indefinite*, due to the indefinite structure of the DeWitt metric G_{abcd}; see the discussion following (4.94). Consequently, the Wheeler–DeWitt equation (5.18), too, possesses an indefinite kinetic term. From this point of view (5.18) resembles a Klein–Gordon equation; strictly speaking, infinitely many Klein–Gordon equations with a non-trivial potential term. This can be made more explicit. Using instead of h_{ab} and $\sqrt{h}$ the variables $\tilde{h}_{ab} = h^{-1/3}h_{ab}$ (cf. (4.98)) and the local volume element $\sqrt{h}$, the Wheeler–DeWitt equation can be written explicitly as

$$\begin{aligned}\Bigg(6\pi G\hbar^2\sqrt{h}\frac{\delta^2}{\delta(\sqrt{h})^2} &- \frac{16\pi G\hbar^2}{\sqrt{h}}\tilde{h}_{ac}\tilde{h}_{bd}\frac{\delta^2}{\delta\tilde{h}_{ab}\delta\tilde{h}_{cd}} \\ &- \frac{\sqrt{h}}{16\pi G}({}^{(3)}R - 2\Lambda)\Bigg)\Psi[\sqrt{h},\tilde{h}_{ab}] = 0.\end{aligned} \tag{5.21}$$

It might therefore be more appropriate to use a *Klein–Gordon-type inner product.* If Ψ_1 and Ψ_2 are solutions of (5.18), the functional version of this inner product would read (DeWitt 1967*a*)

$$\begin{aligned}\langle\Psi_1|\Psi_2\rangle = \mathrm{i}\int\prod_{\mathbf{x}} \mathrm{d}\Sigma^{ab}(\mathbf{x})\Psi_1^*[h_{ab}] \times \\ \left(G_{abcd}\frac{\overrightarrow{\delta}}{\delta h_{cd}} - \frac{\overleftarrow{\delta}}{\delta h_{cd}}G_{abcd}\right)\Psi_2[h_{ab}] = \langle\Psi_2|\Psi_1\rangle^*.\end{aligned} \tag{5.22}$$

Here, the integration is over a $5 \times \infty^3$-dimensional surface in the $6 \times \infty^3$-dimensional space Riem Σ, and $\mathrm{d}\Sigma^{ab}$ denotes the corresponding surface element. In view of (5.21), the integration can be taken over the variables $\tilde{h}_{ab}$, referring to 'constant time $\sqrt{h}$'. Of course, the lack of mathematical rigour is the same as with (5.20).

The inner product (5.22) has the advantage that it is invariant under deformations of the $5 \times \infty^3$-dimensional surface. This expresses its 'time independence'. However, this inner product is—like the usual inner product for the Klein–Gordon equation—not positive definite. In particular, one has $\langle \Psi | \Psi \rangle = 0$ for real solutions of (5.18). Since the Wheeler–DeWitt equation is a real equation (unlike the Schrödinger equation), real solutions should possess some significance.

For the standard Klein–Gordon equation in Minkowski space, one can make a separation between 'positive' frequencies and 'negative' frequencies. As long as one can stay within the one-particle picture, it is consistent to make a restriction to the positive-frequency sector. For such solutions, the inner product is positive. On curved backgrounds, a separation into positive and negative frequencies can be made if both the space–time metric and the potential are stationary, that is, if there is a time-like Killing field and if the potential is constant along its orbits. The Killing field can also be a conformal Killing field, but then the potential must obey certain scaling properties. Moreover, the potential must be positive. If these conditions are violated, particles are produced and the one-particle picture breaks down.

Can such a separation into positive and negative frequencies be made for the Wheeler–DeWitt equation? The clear answer is *no* (Kuchař 1992). There exists a conformal Killing field for the DeWitt metric, namely the three-metric h_{ab}. The potential, however, neither is positive definite nor scales in the correct way. Therefore, no Klein–Gordon inner product can be constructed which is positive definite for the generic case (although this might be achievable for special models). For the standard Klein–Gordon equation, the failure of the one-particle picture leads to 'second quantization' and quantum field theory. The Wheeler–DeWitt equation, however, corresponds already to a field-theoretic situation. It has, therefore, been suggested to proceed with a 'third quantization' and to turn the wave function $\Psi[h]$ into an operator (McGuigan 1988; see also Kuchař 1992 for a review and further references). No final progress, however, has been achieved with such attempts.

One might wonder whether the problems with the above attempts are an indication of the absence of time at the most fundamental level. As will be discussed in Section 5.4, the usual concept of time emerges as an approximate notion at a semiclassical level. This is, in fact, all that is needed to have accordance with experience. Also, the notion of a Hilbert space may be a semiclassical artefact only, because the Hilbert-space structure of standard quantum theory is inextricably linked with the presence of an external time; the reason for this is the probability interpretation, which demands the conservation of probability *in* time (unitarity). If time disappears, the Hilbert space may thus disappear, too.[8] It is, however, not yet clear what kind of mathematical structure could replace it at the fundamental level in order to select physically reasonable solutions from $\mathcal{F}_0$. We shall, therefore, proceed pragmatically and treat

[8]As DeWitt (1999) put it, '...one learns that time and probability are both *phenomenological* concepts.'

the Wheeler–DeWitt equation (5.18) just as a (functional) differential equation in the following.

5.3 The geometrodynamical wave function

5.3.1 The diffeomorphism constraints

The general equations (5.18) and (5.19) are very complicated and need a mathematical elaboration. Even more, the operator appearing in (5.18) needs first to be defined properly. Still, some general features can be studied directly. For the diffeomorphism constraints (5.19), this is much easier to achieve, and the present subsection is therefore devoted to them. Since we have seen in Section 4.2 that the classical constraints $\mathcal{H}_a^{\mathrm{g}}$ generate three-dimensional coordinate transformations, the presence of the quantum constraints (5.19) expresses the invariance of the wave functional Ψ under such transformations or, more precisely, under infinitesimal coordinate transformations.

This can be seen as follows (Higgs 1958). Under the infinitesimal transformation

$$x^a \mapsto \bar{x}^a = x^a + \delta N^a(\mathbf{x}),$$

the three-metric transforms as

$$h_{ab}(\mathbf{x}) \mapsto \bar{h}_{ab}(\mathbf{x}) = h_{ab}(\mathbf{x}) - D_a \delta N_b(\mathbf{x}) - D_b \delta N_a(\mathbf{x}).$$

The wave functional then transforms according to

$$\Psi[h_{ab}] \mapsto \Psi[h_{ab}] - 2 \int \mathrm{d}^3x \; \frac{\delta \Psi}{\delta h_{ab}(\mathbf{x})} D_a \delta N_b(\mathbf{x}).$$

Assuming that $\delta N_b(\mathbf{x})$ vanishes at infinity, one can perform a partial integration and conclude from the arbitrariness of $\delta N_b(\mathbf{x})$ that

$$D_a \frac{\delta \Psi}{\delta h_{ab}} = 0,$$

that is, (5.19) is fulfilled. Therefore, Ψ depends only on the three-dimensional geometry, not on the particular form of the metric, that is, it is implicitly defined on superspace (Section 4.2). This is sometimes expressed by the notation $\Psi[\,^3\mathcal{G}]$ (Wheeler 1968). Such a representation is, however, at best pictorial, since one cannot construct a derivative operator of the form $\delta/\delta(\,^3\mathcal{G})$ on superspace; one must work with the equations (5.18) and (5.19) for $\Psi[h_{ab}]$. This is similar to gauge theories (Section 4.1), where one has to work with the connection and where an explicit transition to gauge-invariant variables is in general impossible. Note that the above formal manipulation to demonstrate coordinate invariance for Ψ is completely analogous to (3.80) for parametrized field theories, and to (3.43) for the bosonic string.

A simple analogy to (5.19) is Gauss's law in QED (or its generalizations to the non-Abelian case; see Section 4.1). The quantized version of the constraint $\nabla \mathbf{E} \approx 0$ reads

$$\frac{\hbar}{\mathrm{i}} \nabla \frac{\delta \Psi[\mathbf{A}]}{\delta \mathbf{A}} = 0, \tag{5.23}$$

from which invariance of Ψ with respect to gauge transformations $\mathbf{A} \to \mathbf{A} + \nabla\lambda$ follows.

We have seen that the wave functional $\Psi[h_{ab}]$ is invariant under infinitesimal coordinate transformations ('small diffeomorphisms'). There may, however, exist 'large diffeomorphisms', that is, diffeomorphisms which are not connected with the identity, under which Ψ might not be invariant.

This situation is familiar from Yang–Mills theories (see e.g. Huang 1992). The quantized form of the Gauss law (4.30) demands that $\Psi[A_a^i]$ be invariant under infinitesimal ('small') gauge transformations; cf. the QED example (5.23). We take the Yang–Mills gauge group $\mathcal{G}$ as the map

$$S^3 \longrightarrow \mathrm{SU}(N) \equiv G, \tag{5.24}$$

where $\mathbb{R}^3$ has been compactified to the three-sphere S^3; this is possible, since it is assumed that gauge transformations approach a constant at spatial infinity. The key role in the study of 'large gauge transformations' is played by

$$\pi_0(\mathcal{G}) \equiv \mathcal{G}/\mathcal{G}_0, \tag{5.25}$$

where $\mathcal{G}_0$ denotes the component of $\mathcal{G}$ connected with the identity. Thus, π_0 counts the number of components of the gauge group. One can also write

$$\pi_0(\mathcal{G}) = [S^3, G] \equiv \pi_3(G) = \mathbb{Z}, \tag{5.26}$$

where $[S^3, G]$ denotes the set of homotopy classes of continuous maps from S^3 to G.[9] The 'winding numbers' $n \in \mathbb{Z}$ denote the number of times that the spatial S^3 is covered by the SU(2) manifold S^3.[10] This then leads to a vacuum state for each connected component of $\mathcal{G}$, called a 'K-vacuum' $|k\rangle$, $k \in \mathbb{Z}$. A state $|k\rangle$ is invariant under small gauge transformations, but transforms as $|k\rangle \to |k+n\rangle$ under large gauge transformations. If one defines the central concept of a 'θ-vacuum' by

$$|\theta\rangle = \sum_{k=-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}k\theta}|k\rangle, \tag{5.27}$$

with a real parameter θ, the transformation of this state under a large gauge transformation reads

$$\sum_{k=-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}k\theta}|k+n\rangle = \mathrm{e}^{\mathrm{i}n\theta}|\theta\rangle.$$

The θ-states are thus labelled by $\mathrm{Hom}(\mathbb{Z}, U(1))$, the homomorphisms from $\mathbb{Z}$ to $U(1)$. Different values of θ characterize different 'worlds' (compare the ambiguity related to the Barbero–Immirzi parameter in Section 4.3); θ is in principle a measurable quantity and one has, for example, from the limit on the neutron dipole moment, the constraint $|\theta| < 10^{-9}$ on the θ-parameter of QCD. Instead of the gauge-dependent wave functions (5.27), one can work with gauge-independent wave functions, but with an additional

[9]Two maps are called homotopic if they can be continuously deformed into each other. All homotopic maps yield a homotopy class.

[10]The SU(N) case can be reduced to the SU(2) case.

term in the action, the 'θ-action' (Ashtekar 1991, Huang 1992). A state of the form (5.27) is also well known from solid state physics (the 'Bloch wave function').

One can envisage the states $|k\rangle$ as being 'peaked' around a particular minimum in a periodic potential. Therefore, tunnelling is possible between different minima. In fact, tunnelling is described by 'instantons', that is, solutions of the classical Euclidean field equations for which the initial and final values of the gauge potential differ by a large gauge transformation (Huang 1992).

One does not have to restrict oneself to S^3, but can generalize this notion to an arbitrary compact orientable three-space Σ (Isham 1981):

$$|\theta\rangle = \sum_{(k,g)} \theta(k,g)|k,g\rangle, \tag{5.28}$$

where $\theta(k,g) \in \mathrm{Hom}([\Sigma, G], U(1))$ appears instead of the $\mathrm{e}^{-\mathrm{i}k\theta}$ of (5.27). As it turns out, $g \in \mathrm{Hom}(\pi_1(\Sigma), \pi_1(G))$.

Instead of taking the gauge group as the starting point, one can alternatively focus on the physical configuration space of the theory. This is more suitable for the comparison with gravity. For Yang–Mills fields, one has the configuration space $Q = \mathcal{A}/\mathcal{G}$, where $\mathcal{A}$ denotes the set of connections on space. In gravity, $Q = \mathcal{S}(\Sigma) = \mathrm{Riem}\ \Sigma/\mathrm{Diff}\ \Sigma$; see Section 4.2. If the group acts freely on $\mathcal{A}$ (or Riem Σ), that is, if it has no fixed points, then

$$\pi_1(Q) = \pi_0(\mathcal{G}),$$

and the θ-structure as obtained from $\pi_0(\mathcal{G})$ can be connected directly with the topological structure of the configuration space, that is, with $\pi_1(Q)$. As we have seen in Section 4.2.5, Diff Σ does not act freely on Riem Σ, so $\mathcal{S}(\Sigma)$ had to be transformed into the 'resolution space' $\mathcal{S}_{\mathrm{R}}(\Sigma)$. Everything is fine if we restrict Diff Σ to $\mathcal{D}_{\mathrm{F}}(\bar{\Sigma})$ (this is relevant in the open case) and take into account the fact that $\mathcal{S}_{\mathrm{R}}(\bar{\Sigma}) \cong \mathcal{S}(\Sigma)$. Then,

$$\pi_1(\mathcal{S}(\Sigma)) = \pi_0(\mathcal{D}_{\mathrm{F}}(\bar{\Sigma})),$$

and one can classify θ-states by elements of $\mathrm{Hom}(\pi_0(\mathcal{D}_{\mathrm{F}}(\bar{\Sigma})), U(1))$. Isham (1981) has investigated the question as to which three-manifolds Σ can yield a non-trivial θ-structure. He found that

$$\pi_0(\mathcal{D}_{\mathrm{F}}(S^3)) = 0,$$

so no θ-structure is available in the cosmologically interesting case S^3. A θ-structure is present, for example, in the case of 'Wheeler's wormhole', that is, for $\Sigma = S^1 \times S^2$. In that case,

$$\pi_0(\mathcal{D}_{\mathrm{F}}(S^1 \times S^2)) = \mathbb{Z}_2 \oplus \mathbb{Z}_2,$$

where $\mathbb{Z}_2 = \{-1, 1\}$. Also, for the three-torus $S^1 \times S^1 \times S^1$, one has a non-vanishing π_0, but the expression is more complicated. Therefore, θ-sectors in quantum gravity are associated with the disconnectedness of Diff Σ. In the asymptotically flat case, something interesting may occur in addition (Friedman and Sorkin 1980). If one allows rotations at infinity, one can get half-integer spin states in the case where a rotation by 2π acts non-trivially, that is, if one cannot communicate a rotation by 2π at ∞ to the whole interior of space. An example of a manifold that allows such states is $\Sigma = \mathbb{R}^3 \# T^3$.

5.3.2 WKB approximation

An important approximation in quantum mechanics is the WKB approximation. On a formal level, this can also be performed for equations (5.18) and (5.19). For this purpose, one makes the ansatz[11]

$$\Psi[h_{ab}] = C[h_{ab}] \exp\left(\frac{\mathrm{i}}{\hbar} S[h_{ab}]\right), \tag{5.29}$$

where $C[h_{ab}]$ is a 'slowly varying amplitude' and $S[h_{ab}]$ is a 'rapidly varying phase' (an 'eikonal' as in geometrical optics). This corresponds to

$$p^{ab} \longrightarrow \frac{\delta S}{\delta h_{ab}},$$

which is the classical relation for the canonical momentum, and from (5.18) and (5.19) one finds the approximate equations

$$16\pi G \, G_{abcd} \frac{\delta S}{\delta h_{ab}} \frac{\delta S}{\delta h_{cd}} - \frac{\sqrt{h}}{16\pi G}({}^{(3)}R - 2\Lambda) = 0, \tag{5.30}$$

$$D_a \frac{\delta S}{\delta h_{ab}} = 0. \tag{5.31}$$

In the presence of matter, one has additional terms. Equation (5.30) is the Hamilton–Jacobi equation for the gravitational field (Peres 1962). Equation (5.31) expresses again the fact that $S[h_{ab}]$ is invariant under coordinate transformations. One can show that (5.30) and (5.31) are fully equivalent to the classical Einstein field equations (Gerlach 1969)—this is one of the 'six routes to geometrodynamics' (Misner *et al.* 1973). This route again shows how the dynamical laws follow from the laws of the instant (Kuchař 1993).

The 'interconnection theorems' mentioned in Section 4.2 have their counterparts on the level of the eikonal $S[h_{ab}]$. For example, if S satisfies (5.30), it must automatically satisfy (5.31). These relations have their counterpart in the full quantum theory, provided there are no anomalies (Section 5.3.5).

More useful than a WKB approximation for all degrees of freedom is a 'mixed' approximation scheme in which gravity is treated differently from other fields. It is then possible to recover the limit of quantum field theory in an external space–time. This method will be presented in Section 5.4.

Another approximation that might be applied to the Wheeler–DeWitt equation (5.18) is the strong-coupling limit (Isham 1976). This is the quantum analogue of the limit $G \to \infty$ mentioned at the end of Section 4.2.3. Using the ad hoc regularization that $\delta(0)$-terms vanish, one can find simple exponential functionals that solve (5.18) and (5.19). The exponent is then inversely proportional to the square root of minus the cosmological constant. For positive Λ, one thus obtains an oscillatory exponential, while for negative Λ the exponential is real.

[11]In contrast to standard quantum theory, the occurrence here of the imaginary unit i is somewhat artificial because the constraint equations for Ψ are real (Barbour 1993).

5.3.3 Remarks on the functional Schrödinger picture

The central kinematical object in quantum geometrodynamics is the wave functional. It obeys the quantum constraint equations (5.18) and (5.19), which are functional differential equations. In non-gravitational quantum field theory, this 'Schrödinger picture' is used only infrequently, mainly because the focus there is on perturbative approaches for which other formulations are more appropriate, for example, the Fock-space picture. Still, even there a Schrödinger picture is sometimes used in the form of the 'Tomonaga–Schwinger' (TS) equation (known as such although it was first formulated by Stueckelberg 1938),

$$\mathrm{i}\hbar\frac{\delta\Psi}{\delta\tau(\mathbf{x})} = \mathcal{H}\Psi, \tag{5.32}$$

where $\tau(\mathbf{x})$ is the local 'bubble time' parameter, and $\mathcal{H}$ is the Hamiltonian density; for instance, in the case of a scalar field, one has

$$\mathcal{H} = -\frac{\hbar^2}{2}\frac{\delta^2}{\delta\phi^2} + \frac{1}{2}(\nabla\phi)^2 + V(\phi). \tag{5.33}$$

In the 'choice of time before quantization' approach, the gravitational constraints are directly cast into TS form, see (5.14). Equation (5.32) is at best only of formal significance. First, the bubble time cannot be a scalar; see Giulini and Kiefer (1995). Second, although the TS equation describes in principle the evolution along all possible foliations of space–time into space-like hypersurfaces, this evolution cannot be unitarily implemented on Fock space (Giulini and Kiefer 1995, Helfer 1996, Torre and Varadarajan 1999). The only sensible approach is the use of a genuine Schrödinger equation, that is, an integrated version of (5.32) along a privileged foliation of space–time (such as the one given by York's time (5.16)).

The Schrödinger equation can be applied successfully to some non-perturbative aspects of quantum field theory; see Jackiw (1995), Section IV.4, for a detailed discussion and references. Among the applications are the θ-structure of QCD (cf. Section 5.3.1), chiral anomalies, and confinement. From a more fundamental point of view one can show that, at least for the ϕ^4-theory in Minkowski space, the Schrödinger picture exists at each order of perturbation theory; that is, in each order, one has an integrated version of (5.32) with renormalized quantities,

$$\mathrm{i}\hbar\frac{\partial\Psi_{\mathrm{ren}}}{\partial t} = \int \mathrm{d}^3x\ \mathcal{H}_{\mathrm{ren}}\Psi_{\mathrm{ren}}, \tag{5.34}$$

where one additional renormalization constant is needed in comparison to the Fock-space formulation; see Symanzik (1981).[12]

[12] This analysis was generalized by McAvity and Osborn (1993) to quantum field theory on manifolds with arbitrarily smoothly curved boundaries. Non-Abelian fields are treated, for example, in Lüscher *et al.* (1992).

The simplest example of the Schrödinger picture is provided by the free bosonic field. The implementation of the commutation relations

$$[\hat{\phi}(\mathbf{x}), \hat{p}_\phi(\mathbf{y})] = \mathrm{i}\hbar\delta(\mathbf{x} - \mathbf{y}) \tag{5.35}$$

leads to

$$\hat{\phi}(\mathbf{x})\Psi[\phi(\mathbf{x})] = \phi(\mathbf{x})\Psi[\phi(\mathbf{x})], \tag{5.36}$$

$$\hat{p}_\phi\Psi[\phi(\mathbf{x})] = \frac{\hbar}{\mathrm{i}}\frac{\delta}{\delta\phi(\mathbf{x})}\Psi[\phi(\mathbf{x})], \tag{5.37}$$

where $\Psi[\phi(\mathbf{x})]$ is a wave functional on the space of all fields $\phi(\mathbf{x})$, which includes not only smooth classical configurations but also distributional ones. The Hamilton operator for a free massive scalar field reads (from now on again, $\hbar = 1$)

$$\begin{aligned}\hat{H} &= \frac{1}{2}\int \mathrm{d}^3x\ \left(\hat{p}_\phi^2(\mathbf{x}) + \hat{\phi}(\mathbf{x})(-\nabla^2 + m^2)\hat{\phi}(\mathbf{x})\right)\\ &= \frac{1}{2}\int \mathrm{d}^3x\ \hat{p}_\phi^2(\mathbf{x}) + \frac{1}{2}\int \mathrm{d}^3x\,\mathrm{d}^3x'\ \hat{\phi}(\mathbf{x})\omega^2(\mathbf{x},\mathbf{x}')\hat{\phi}(\mathbf{x}'),\end{aligned} \tag{5.38}$$

where

$$\omega^2(\mathbf{x},\mathbf{x}') := (-\nabla^2 + m^2)\delta(\mathbf{x} - \mathbf{x}') \tag{5.39}$$

is not diagonal in three-dimensional space, but is diagonal in momentum space (due to translation invariance):

$$\begin{aligned}\omega^2(\mathbf{p},\mathbf{p}') &:= \int \mathrm{d}^3p''\ \omega(\mathbf{p},\mathbf{p}'')\omega(\mathbf{p}'',\mathbf{p}')\\ &= \frac{1}{(2\pi)^3}\int \mathrm{d}^3x\,\mathrm{d}^3x'\ \mathrm{e}^{\mathrm{i}\mathbf{px}}\omega^2(\mathbf{x},\mathbf{x}')\mathrm{e}^{-\mathrm{i}\mathbf{p}'\mathbf{x}'}\\ &= (p^2 + m^2)\delta(\mathbf{p} - \mathbf{p}'),\end{aligned} \tag{5.40}$$

with $p := |\mathbf{p}|$. Therefore,

$$\omega(\mathbf{p},\mathbf{p}') = \sqrt{p^2 + m^2}\,\delta(\mathbf{p} - \mathbf{p}') := \omega(p)\,\delta(\mathbf{p} - \mathbf{p}'). \tag{5.41}$$

The stationary Schrödinger equation then reads (we set $\hbar = 1$)

$$\hat{H}\Psi_n[\phi] \equiv \left(-\frac{1}{2}\int \mathrm{d}^3x\ \frac{\delta^2}{\delta\phi^2} + \frac{1}{2}\int \mathrm{d}^3x\,\mathrm{d}^3x'\ \phi\omega^2\phi\right)\Psi_n[\phi] = E_n\Psi_n[\phi]. \tag{5.42}$$

In analogy to the ground-state wave function of the quantum-mechanical harmonic oscillator, the ground-state solution of (5.42) reads

$$\Psi_0[\phi] = \det{}^{1/4}\left(\frac{\omega}{\pi}\right)\exp\left(-\frac{1}{2}\int \mathrm{d}^3x\,\mathrm{d}^3x'\ \phi(\mathbf{x})\omega(\mathbf{x},\mathbf{x}')\phi(\mathbf{x}')\right), \tag{5.43}$$

with the ground-state energy given by

$$E_0 = \frac{1}{2}\mathrm{tr}\ \omega = \frac{1}{2}\int \mathrm{d}^3x\,\mathrm{d}^3x'\ \omega(\mathbf{x},\mathbf{x}')\delta(\mathbf{x} - \mathbf{x}') = \frac{1}{2}\frac{V}{(2\pi)^3}\int \mathrm{d}^3p\ \omega(p). \tag{5.44}$$

This is just the sum of the ground-state energies of infinitely many harmonic oscillators. Not surprisingly, it contains divergences: an infrared (IR) divergence connected with the spatial volume V (due to translational invariance) and an ultraviolet (UV) divergence connected with the sum over all oscillators. This is the usual field-theoretic divergence of the ground-state energy and can be dealt with by standard methods (e.g. normal ordering). Note that the normalization factor in (5.43) is also divergent:

$$\det{}^{1/4}\left(\frac{\omega}{\pi}\right) = \exp\left(\frac{1}{4}\mathrm{tr}\ln\frac{\omega}{\pi}\right) = \exp\left(\frac{V}{32\pi^3}\int \mathrm{d}^3p\ \ln\frac{\sqrt{p^2+m^2}}{\pi}\right). \tag{5.45}$$

One can define the many-particle states (the Fock space) in the usual manner through the application of creation operators on (5.43). The divergence (5.45) cancels in matrix elements between states in the Fock space. However, the space of wave functionals is much bigger than the Fock space. In fact, because there is no unique ground state in the case of time-dependent external fields, any Gaussian functional is called a 'vacuum state', independent of whether it is the ground state of some Hamiltonian or not. A general Gaussian is of the form

$$\Psi_\Omega[\phi] = \det{}^{1/4}\left(\frac{\Omega_R}{\pi}\right)\exp\left(-\frac{1}{2}\int \mathrm{d}^3x\,\mathrm{d}^3x'\ \phi(\mathbf{x})\Omega(\mathbf{x},\mathbf{x}')\phi(\mathbf{x}')\right), \tag{5.46}$$

where $\Omega =: \Omega_R + \mathrm{i}\Omega_I$ is in general complex and time-dependent. One can define in the usual manner an annihilation operator (omitting the integration variables for simplicity)

$$A = \frac{1}{\sqrt{2}}\int \Omega_R^{-1/2}\left(\Omega\phi + \frac{\delta}{\delta\phi}\right) \tag{5.47}$$

and a creation operator

$$A^\dagger = \frac{1}{\sqrt{2}}\int \Omega_R^{-1/2}\left(\Omega^*\phi - \frac{\delta}{\delta\phi}\right). \tag{5.48}$$

One has the usual commutation relation $[A(\mathbf{x}), A^\dagger(\mathbf{y})] = \delta(\mathbf{x}-\mathbf{y})$, and the vacuum state (5.46) is annihilated by A, that is, $A\Psi_\Omega = 0$.

Gaussian functionals are used frequently in quantum field theory with external fields. Examples are an external electric field in QED and an external de Sitter-space background in a gravitational context (Jackiw 1995). The functional Schrödinger picture can also be formulated for fermions (see e.g. Kiefer and Wipf 1994, Jackiw 1995, Barvinsky *et al.* 1999*b*). In the context of linearized gravity, we have already encountered the Gaussian functional describing the graviton ground state; see the end of Section 2.1. Many discussions of the geometrodynamical wave functional take their inspiration from the properties of the Schrödinger picture discussed above.

5.3.4 Connection with path integrals

We have discussed in Section 2.2 the formulation of a quantum-gravitational path integral. In quantum mechanics, the path integral can be shown to satisfy the Schrödinger equation (Feynman and Hibbs 1965). It is therefore of interest to see if a similar property holds in quantum gravity, that is, if the quantum-gravitational path integral (2.72) obeys the quantum constraints (5.18) and (5.19). This is not straightforward, since there are two major differences from ordinary quantum theory. First, one has constraints instead of the usual Schrödinger equation. Second, the path integral (2.72) contains an integration over the whole four-metric, that is, including 'time' (in the form of the lapse function). Since the ordinary path integral in quantum mechanics is a propagator, denoted by $\langle q'', T|q', 0\rangle$, the quantum-gravitational path integral corresponds to an expression of the form

$$\int \mathrm{d}T\ \langle q'', T|q', 0\rangle = G(q'', q'; E)|_{E=0},$$

where the 'energy Green function'

$$G(q'', q'; E) := \int \mathrm{d}T\ \mathrm{e}^{\mathrm{i}ET} \langle q'', T|q', 0\rangle \tag{5.49}$$

has been introduced. The quantum-gravitational path integral thus resembles an energy Green function instead of a propagator and, due to the T-integration, no composition law in the sense of (2.68) holds (Kiefer 1991, 2001*a*). All this is, of course, already true for the models with reparametrization invariance discussed in Section 3.1. In general, an integration over T yields a divergence. One therefore has to choose appropriate contours in a complex T-plane in order to get a sensible result.

A formal derivation of the constraints from (2.72) is straightforward (Hartle and Hawking 1983). Taking a matter field ϕ into account, the path integral reads

$$Z = \int \mathcal{D}g\mathcal{D}\phi\ \mathrm{e}^{\mathrm{i}S[g,\phi]}, \tag{5.50}$$

where the integration over $\mathcal{D}g$ includes an integration over the three-metric, as well as the lapse function N and the shift vector N^a. From the demand that Z be independent of N and N^a at the three-dimensional boundaries, one gets

$$\frac{\delta Z}{\delta N} = 0 = \int \mathcal{D}g\mathcal{D}\phi\ \left.\frac{\delta S}{\delta N}\right|_{\Sigma} \mathrm{e}^{\mathrm{i}S[g,\phi]},$$

where Σ stands for the three-dimensional boundaries, and one gets an analogous expression for N^a. The conditions that the path integrals containing $\delta S/\delta N$ and $\delta S/\delta N^a$ vanish immediately yield the constraints (5.18) and (5.19).

A more careful derivation has to take care of the definition of the measure in the path integral. This was first attempted by Leutwyler (1964), but without taking ghost terms into account. Regarding the correct gauge-fixing procedure, this was achieved by Barvinsky and collaborators; see Barvinsky (1993*a*) for a review and references. Halliwell and Hartle (1991) addressed general reparametrization-invariant systems and demanded that the 'sum over histories' in the path integral respect the invariance generated by the constraints.[13] They assume a set of constraints $H_\alpha \approx 0$ obeying the Poisson-bracket relations

[13]Restriction is made to quantum-mechanical systems, so issues such as field-theoretic anomalies are not discussed.

$$\{H_\alpha, H_\beta\} = U^\gamma_{\alpha\beta} H_\gamma, \tag{5.51}$$

where the $U^\gamma_{\alpha\beta}$ may depend on the canonical variables p_i and q^i (as happens in GR, where a dependence on the three-metric is present; see (3.84)). The corresponding action is written in the form

$$S[p_i, q^i, N^\alpha] = \int_{t_1}^{t_2} \mathrm{d}t\ (p_i \dot{q}^i - N^\alpha H_\alpha), \tag{5.52}$$

where the N^α are Lagrange parameters. As in Section 3.1, one considers

$$\delta p_i = \{p_i, \epsilon^\alpha H_\alpha\}\ , \quad \delta q^i = \{q^i, \epsilon^\alpha H_\alpha\}.$$

If

$$\delta N^\alpha = \dot{\epsilon}^\alpha - U^\alpha_{\beta\gamma} N^\beta \epsilon^\gamma,$$

the action transforms as

$$\delta S = \left[\epsilon^\alpha F_\alpha(p_i, q^i)\right]_{t_1}^{t_2}\ , \quad F_\alpha = p_i \frac{\partial H_\alpha}{\partial p_i} - H_\alpha.$$

Except for constraints linear in the momenta, the action is only invariant if $\epsilon^\alpha(t_2) = 0 = \epsilon^\alpha(t_1)$; cf. (3.28). Halliwell and Hartle (1991) showed that for such systems—together with five natural assumptions about the path integral—the quantum constraints

$$\hat{H}_\alpha \psi(q^i) = 0$$

follow from the path integral. At least at the formal level (neglecting anomalies etc.), GR is included as a special case and the derivation applies. The constraints follow only if the integration range $-\infty < N < \infty$ holds for the lapse function. A *direct* check that the path integral solves the quantum constraints was achieved by Barvinsky (1998) for generic (first-class) constrained systems at the one-loop level of the semiclassical approximation; see also Barvinsky (1993*a*) and the references therein.

The connection between the canonical and the covariant approach can thus be formulated as follows. Under the assumption that all quantities can be rigorously defined, the *full* quantum-gravitational path integral with the action given only by the Einstein–Hilbert action should solve the *full* Wheeler–DeWitt equation and the *full* momentum constraints. In a semiclassical expansion, one can approximate the path integral by an expression containing an effective action with higher curvature terms (see Chapter 2). This does *not*, however, mean that one has to include these higher-order terms in the canonical formalism; at a given order of the semiclassical approximation, the solution of the path integral, which contains the effective action, should coincide with the corresponding solution of the quantum constraints.

Quantum-gravitational path integrals also play a crucial role in the formulation of boundary conditions in quantum cosmology; cf. Section 8.3.

5.3.5 Anomalies and factor ordering

If classical constraints $\mathcal{G}_a \approx 0$ are quantized à la Dirac, one gets a restriction on wave functions according to $\hat{\mathcal{G}}_a \psi = 0$. Also, it is evident that the commutator between two constraints must vanish if applied on wave functions, that is,

$$[\hat{\mathcal{G}}_a, \hat{\mathcal{G}}_b]\psi = 0. \tag{5.53}$$

This requirement is known as 'Dirac consistency'. It only holds if the commutator has the form

$$[\hat{\mathcal{G}}_a, \hat{\mathcal{G}}_b]\psi = C^c_{ab}(\hat{p}, \hat{q})\hat{\mathcal{G}}_c \psi, \tag{5.54}$$

with the coefficients $C^c_{ab}(\hat{p}, \hat{q})$ standing to the *left* of the constraints. If this is not the case, additional terms proportional to a power of $\hbar$ appear. If they are c-numbers as opposed to q-numbers, they are called 'central terms'. More generally, one speaks of 'Schwinger terms' or simply *anomalies.* We have encountered anomalies already in our discussion of the bosonic string; cf. (3.56). If there are anomalies, it is not possible to implement all constraints in the quantum theory via the equations $\hat{\mathcal{G}}_a \psi = 0$.

The question whether there are anomalies in quantum gravity has not yet been answered. One may hope that the demand for an absence of anomalies may fix the factor ordering of the theory and perhaps other issues such as the allowed number of space–time dimensions or the value of fundamental parameters. The question is: can the structure of the Poisson algebra (3.84)–(3.86) be preserved for the corresponding commutators? Or are there necessarily anomalous terms? In spite of much literature, no definite result has arisen. Moreover, Dirac (1968) and Tsamis and Woodard (1987) have emphasized that in order to establish Dirac consistency, one must first properly regularize singular operator products. Otherwise, one can get any result. This is because identities such as

$$f(y)g(x)\delta(x-y) = f(x)g(x)\delta(x-y)$$

are justified for test functions f and g, but not for distributions (and field operators are distributions!).

It might be possible to get some insight by looking at anomalies in ordinary quantum field theories (see e.g. Jackiw 1995 and Bertlmann 1996). In Yang–Mills theories, one has the generalized Gauss law (see (4.30) and (4.124))

$$\mathcal{G}_i := \mathcal{D}_a E^a_i - \rho_i \approx 0, \tag{5.55}$$

where ρ_i denotes the (non-Abelian) charge density for the sources. Quantization yields $\hat{\mathcal{G}}_i \Psi = 0$, where Ψ depends on the gauge fields $A^i_a(x)$ and the charged fields. One also uses instead of A^i_a the variable $\mathcal{A}_a = A^i_a T_i$, where the T_i denote the generators of the gauge group,

$$[T_i, T_j] = C^k_{ij} T_k.$$

Classically,

$$\{\mathcal{G}_i(x), \mathcal{G}_j(y)\} = C^k_{ij} \mathcal{G}_k(x)\delta(x,y). \tag{5.56}$$

Dirac consistency would then be implemented in the quantum theory if one had

$$[\hat{\mathcal{G}}_i(x), \hat{\mathcal{G}}_j(y)] = \mathrm{i}\hbar C_{ij}^k \hat{\mathcal{G}}_k(x)\delta(x,y). \tag{5.57}$$

It is, however, known that anomalies may occur in the presence of fermions with definite chirality (cf. Bertlmann (1996) and the references therein):

$$[\hat{\mathcal{G}}_i(x), \hat{\mathcal{G}}_j(y)] = \mathrm{i}\hbar C_{ij}^k \hat{\mathcal{G}}_k(x)\delta(x,y) \pm \frac{\mathrm{i}\hbar}{24\pi^2}\epsilon^{abc}\mathrm{tr}\{T_i, T_j\}\partial_a \mathcal{A}_b \partial_c \delta(x,y), \tag{5.58}$$

where the sign in front of the second term on the right-hand side depends on the chirality. In perturbation theory, the occurrence of such anomalies arises through triangle graphs. An anomaly is harmless as long as it describes only the breakdown of an external symmetry in the presence of gauge fields. It can then even be responsible for particle decays: the decay $\pi^0 \to \gamma\gamma$, for instance, is fully generated by the axial anomaly (the non-conservation of the axial current). An analogous anomaly, which is of relevance for gravity, is the 'Weyl anomaly' or 'trace anomaly' (see e.g. DeWitt 1979, 2003 and Birrell and Davies 1982). Compare also Section 2.2.4: the invariance of a classical action under Weyl transformations (multiplication of the metric by a function) leads to a traceless energy–momentum tensor. Upon quantization, however, the trace can pick up a non-vanishing term proportional to $\hbar$.

An anomaly becomes problematic if the gauge fields are treated as quantized internal fields because this would lead to a violation of gauge invariance, with all its consequences, such as the destruction of renormalizability. This is the situation described by (5.58). In the Standard Model of strong and electroweak interactions, such harmful anomalies could in principle emerge from the electroweak ($\mathrm{SU}(2) \times \mathrm{U}(1)$) sector. However, the respective anomalies of quarks and leptons cancel each other to render the Standard Model anomaly-free. This is only possible because there are equal numbers of quarks and leptons and because quarks have three possible colours. In superstring theory (Chapter 9), anomaly cancellation occurs (for the heterotic string) only if the possible gauge groups are strongly constrained (either SO(32) or $\mathrm{E8} \times \mathrm{E8}$). Chiral fermions in external gravitational fields lead to (Lorentz) anomalies only in space–time dimensions $2, 6, 10, \ldots$; see, for example, Leutwyler (1986).

It is possible to quantize an anomalous theory, but not through the equations $\hat{\mathcal{G}}_i\Psi = 0$ (Bertlmann 1996). The chiral Schwinger model (chiral QED in 1+1 dimensions), for example, allows a consistent quantum theory with a massive boson.

Another instructive example for the discussion of anomalies is dilaton gravity in 1+1 dimensions. It is well known that GR in 1+1 dimensions possesses no dynamics (see e.g. Brown (1988) for a review), since

$$\int \mathrm{d}^2x\, \sqrt{-g}\, {}^{(2)}R = 4\pi\chi, \tag{5.59}$$

where ${}^{(2)}R$ is the two-dimensional Ricci scalar, and χ is the Euler characteristic of the two-dimensional manifold; if the manifold were a closed compact Riemann surface with genus $\mathfrak{g}$, one would have $\chi = 2(1 - \mathfrak{g})$. Although (5.59) plays a role in string perturbation theory (see Chapter 9), it is of no use in a direct quantization of GR. One can, however, construct non-trivial models in two dimensions if there are degrees of freedom in the gravitational sector in addition to the metric. A particular example is the presence of a *dilaton field*. Such a field occurs, for example, in the 'CGHS model' presented in Callan *et al.* (1992). This model is defined by the action

$$4\pi G S_{\mathrm{CGHS}} = \int \mathrm{d}^2x\,\sqrt{-g}\;\mathrm{e}^{-2\phi}\left({}^{(2)}R + 4g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \lambda\right) + S_{\mathrm{m}}, \tag{5.60}$$

where ϕ is the dilaton field, and λ is a parameter ('cosmological constant') with dimension L^{-2}.[14] Note that the gravitational constant G is dimensionless in two dimensions. The name 'dilaton' comes from the fact that ϕ occurs in the combination $\mathrm{d}^2x\sqrt{-g}\,\mathrm{e}^{-2\phi}$ and can thus be interpreted as describing an effective change of integration measure (a 'change of volume'). It is commonly found in string perturbation theory (Chapter 9), and its value there determines the string coupling constant.

The simplest choice for the matter action S_{m} is an ordinary scalar-field action,

$$S_{\mathrm{m}} = \frac{1}{2}\int \mathrm{d}^2x\;\sqrt{-g}\,g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi. \tag{5.61}$$

Cangemi *et al.* (1996) make a series of redefinitions and canonical transformations (partly non-local) to simplify this action. The result is then defined as providing the starting point for quantization (independent of whether equivalence to the old variables holds or not). In the Hamiltonian version, one finds constraints again: one Hamiltonian constraint and one momentum constraint. They read (after a rescaling $\lambda \to \lambda/8\pi G$)

$$\mathcal{H}_\perp = \frac{(\pi_1)^2 - (\pi_0)^2}{2\lambda} - \frac{\lambda}{2}([r^0]')^2 + \frac{\lambda}{2}([r^1]')^2 + \frac{1}{2}(\pi_\varphi^2 + [\varphi']^2), \tag{5.62}$$

$$\mathcal{H}_1 = -[r^0]'\pi_0 - [r^1]'\pi_1 - \varphi'\pi_\varphi, \tag{5.63}$$

where r^0 and r^1 denote the new gravitational variables (found from the metric—the only dynamical part being its conformal part—and the dilaton), and π_0 and π_1 are their respective momenta. The form of $\mathcal{H}_1$ is similar to that in the case of parametrized field theory and string theory; cf. (3.74) and (3.43). One recognizes explicitly that the kinetic term in $\mathcal{H}_\perp$ is indefinite. In fact, the Hamiltonian constraint describes an 'indefinite harmonic oscillator' (Zeh 1988)—the sum of two ordinary oscillators with opposite signs (see also Section 8.1.2).

According to our general prescription, one has in the quantum theory

$$\hat{\mathcal{H}}_\perp\Psi(r^0, r^1, \varphi) = 0\ , \quad \hat{\mathcal{H}}_1\Psi(r^0, r^1, \varphi) = 0. \tag{5.64}$$

Although both $\mathcal{H}_\perp$ and $\mathcal{H}_1$ are sums of independent terms, one cannot expect to find a product state with respect to r^0, r^1, and φ as a common solution to the two equations. The resulting quantum state will be entangled. (This is called 'correlation interaction' by Cangemi *et al.* (1996)).

[14] One can exhaust all dilaton models by choosing instead of λ any potential $V(\phi)$; cf. Louis-Martinez and Kunstatter (1994). A particular example is the dimensional reduction of spherically symmetric gravity to two dimensions; see Grumiller *et al.* (2002) for a general review of dilaton gravity in two dimensions.

The algebra of constraints (3.84)–(3.86) in the quantum theory reads[15]

$$\mathrm{i}[\hat{\mathcal{H}}_\perp(x), \hat{\mathcal{H}}_\perp(y)] = \hbar(\hat{\mathcal{H}}_1(x) + \hat{\mathcal{H}}_1(y))\delta'(x-y), \tag{5.65}$$

$$\mathrm{i}[\hat{\mathcal{H}}_\perp(x), \hat{\mathcal{H}}_1(y)] = \hbar(\hat{\mathcal{H}}_\perp(x) + \hat{\mathcal{H}}_\perp(y))\delta'(x-y) - \frac{c\hbar^2}{12\pi}\delta'''(x-y), \tag{5.66}$$

$$\mathrm{i}[\hat{\mathcal{H}}_1(x), \hat{\mathcal{H}}_1(y)] = \hbar(\hat{\mathcal{H}}_1(x) + \hat{\mathcal{H}}_1(y))\delta'(x-y). \tag{5.67}$$

Note the absence of the metric on the right-hand side of these equations. This is different from the (3+1)-dimensional case. The reason is that $hh^{ab} = 1$ in one spatial dimension and that the constraint generators have been rescaled by a factor $\sqrt{h}$.

In (5.66), an additional 'Schwinger term' with central charge c has been added. The reason is a theorem by Boulware and Deser (1967), stating that there must necessarily be a Schwinger term in the commutator

$$[\hat{\mathcal{H}}_\perp(x), \hat{\mathcal{H}}_1(y)].$$

This theorem was proven, however, within standard Poincaré-invariant local field theory, with the additional assumption that there is a ground state of the Hamiltonian. This is certainly not a framework that is applicable in a gravitational context. But since the equations (5.62) and (5.63) have the form of equations in flat space–time, one can tentatively apply this theorem. The central charge is then a sum of three contributions (Cangemi *et al.* 1996),

$$c = c^{\mathrm{g}} + c^{\mathrm{m}} \equiv c_0^{\mathrm{g}} + c_1^{\mathrm{g}} + c^{\mathrm{m}}, \tag{5.68}$$

where c_0^{g} and c_1^{g} are the central charges connected with the gravitational variables r^0 and r^1, respectively, and c^{m} is the central charge connected with the field φ. The result for c depends on the notion of a vacuum (if there is one). Standard methods (decomposition into creation and annihilation operators) yield $c_1^{\mathrm{g}} = 1$. What can be said about c_0^{g}? If the sign in front of the $(\pi_0)^2$ term in (5.62) were positive, one would have $c_0^{\mathrm{g}} = 1$ too. But with the minus sign, one cannot simultaneously demand a positive energy and a positive norm. If one demands a positive norm, one must associate a positive frequency with the creation operator (instead of, as usual, the annihilation operator). This would yield $c_0^{\mathrm{g}} = -1$. Then, $c = 0$ in the absence of the φ-field and one would have no anomaly. The constraints (5.64) can then be consistently imposed, and one can find the following two solutions in the pure gravitational case:

$$\Psi_{\mathrm{g}}(r^0, r^1) = \exp\left(\pm\frac{\mathrm{i}\lambda}{2\hbar}\int \mathrm{d}x\ (r^0[r^1]' - r^1[r^0]')\right). \tag{5.69}$$

Exact states describing black holes in generic dilaton models are discussed in Barvinsky and Kunstatter (1996).

The presence of the φ-field would, however, yield $c^{\mathrm{m}} = 1$ and the anomaly would not vanish; $c = 1$. Cangemi *et al.* (1996) have shown that the anomaly can be cancelled by adding an appropriate counterterm, but this leads to a complicated form of the

[15] The Virasoro form (3.56) of the algebra follows for the combinations $\theta_\pm = (\mathcal{H}_\perp \mp \mathcal{H}_1)$.

quantum constraints for which no solution is in sight.[16] Albeit obtained within an unrealistically simple model, the above discussion demonstrates what kind of problems can be expected to occur. The presence of anomalies might prevent one from imposing all constraints in GR à la Dirac, but one could also imagine that in the full theory a cancellation of the various central charges might occur. The latter is suggested by the indefinite kinetic term in quantum gravity, but an explicit demonstration is far from reach.

The above discussion of anomalies refers to a field-theoretic context. However, even for finite-dimensional models with constraints $\mathcal{H}_\perp \approx 0$, $\mathcal{H}_a \approx 0$, the demand for closure of the quantum algebra leads to restrictions on the possible factor ordering. This was studied by Barvinsky and Krykhtin (1993) for general constrained systems and applied by Barvinsky (1993*b*) to the gravitational case. The classical constraints are again collectively written as $\mathcal{H}_\alpha \approx 0$. Their Poisson-bracket algebra—the analogue to (3.84)–(3.86)—is given by the shorthand notation (5.51). Demanding equivalence of 'Dirac quantization' and 'BRST quantization', Barvinsky and Krykhtin (1993) find the relation

$$\hat{\mathcal{H}}_\alpha - \hat{\mathcal{H}}_\alpha^\dagger = \mathrm{i}\hbar \left(\hat{U}^\lambda_{\alpha\lambda}\right)^\dagger + \mathcal{O}(\hbar^2), \tag{5.70}$$

where the adjoint is defined with respect to the standard Schrödinger inner product. The constraints are thus not self-adjoint, which anyway is expected (Komar 1979). If one demands that the constraints be covariant with respect to redefinitions in configuration space, their quantum form is fixed to read (with $32\pi G = 1$)

$$\hat{\mathcal{H}}_\perp = -\frac{\hbar^2}{2} G_{abcd} \mathcal{D}^{ab} \mathcal{D}^{cd} + V, \tag{5.71}$$

$$\hat{\mathcal{H}}_a = -\frac{2\hbar}{\mathrm{i}} D_b h_{ac} \mathcal{D}^{bc} + \frac{\mathrm{i}\hbar}{2} U^\lambda_{a\lambda}, \tag{5.72}$$

where $\mathcal{D}^{ab}$ is the covariant derivative with respect to the DeWitt metric,

$$\mathcal{D}^{ab} := \frac{\mathcal{D}\Psi}{\mathcal{D}h_{ab}}.$$

The big open problem is of course to see whether this result survives the transition to the field-theoretic case, that is, whether no anomalies are present after a consistent regularization has been performed.

5.3.6 Canonical quantum supergravity

We have seen in Section 2.3 that the quantization of *super*gravity instead of GR exhibits interesting features. Thus, it seems worthwhile to discuss the canonical quantization of supergravity (SUGRA). This will be briefly reviewed here; more details can be found in D'Eath (1984, 1996) and Moniz (2010). The situation in three space–time dimensions is addressed in Nicolai and Matschull (1993). Here, we will only consider the case of $N = 1$ SUGRA given by the action (2.143).

[16] These authors also present a proposal for BRST quantization, $Q_{\mathrm{BRST}}\Psi = 0$, where Q_{BRST} is the BRST charge; cf. Section 9.1. They find many solutions, which, however, cannot be properly interpreted.

The classical canonical formalism was developed by Fradkin and Vasiliev (1977), Pilati (1977), and Teitelboim (1977). Working again with tetrads (cf. Section 1.1), one has

$$g_{\mu\nu} = \eta_{nm} e^n_\mu e^m_\nu . \tag{5.73}$$

For the quantization, it is more convenient to use two-component spinors according to

$$e^{AA'}_\mu = e^n_\mu \sigma^{AA'}_n , \tag{5.74}$$

where A runs from 1 to 2, A' runs from $1'$ to $2'$, and the van der Waerden symbols $\sigma^{AA'}_n$ denote the components of the matrices

$$\sigma_0 = -\frac{1}{\sqrt{2}}\mathbb{I}\ , \quad \sigma_a = \frac{1}{\sqrt{2}} \times \text{Pauli matrix}, \tag{5.75}$$

with raising and lowering of indices by ϵ^{AB}, ϵ_{AB}, $\epsilon^{A'B'}$, $\epsilon_{A'B'}$, which are all given in matrix form by

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} ;$$

see, for example, Wess and Bagger (1992) and Sexl and Urbantke (2001) for more details of this formalism. The inverse of (5.74) is given by

$$e^n_\mu = -\sigma^n_{AA'} e^{AA'}_\mu , \tag{5.76}$$

where $\sigma^n_{AA'}$ is obtained from $\sigma^{AA'}_n$ by raising and lowering indices. One can go from tensors to spinors via $e^{AA'}_\mu$ and from spinors to tensors via $e^\mu_{AA'}$.

One can now rewrite the action (2.143) in two-component language. Instead of e^n_μ (vierbein) and ψ^α_μ (gravitino), one works with the spinor-valued one-form $e^{AA'}_\mu$ and the spinor-valued one-form ψ^A_μ plus its Hermitian conjugate $\bar{\psi}^{A'}_\mu$. The latter two are odd Grassmann variables; that is, they are anticommuting among themselves. The action then reads (for $\Lambda = 0$)

$$S = \frac{1}{16\pi G}\int \mathrm{d}^4x\ (\det e^n_\mu) R + \frac{1}{2}\int \mathrm{d}^4x\ \epsilon^{\mu\nu\rho\sigma}\left(\bar{\psi}^{A'}_\mu e_{AA'\nu} D_\rho \psi^A_\sigma + \text{h.c.}\right) . \tag{5.77}$$

The derivative D_ρ acts on spinor-valued forms (i.e. it acts on their spinor indices only),

$$D_\mu \psi^A_\nu = \partial_\mu \psi^A_\nu + \omega^A_{B\mu}\psi^B_\nu , \tag{5.78}$$

where $\omega^A_{B\mu}$ denotes the spinorial version of ω^{nm}_μ (see D'Eath 1984). We remark that the presence of gravitinos leads to torsion,

$$D_{[\mu} e^{AA'}_{\nu]} = S^{AA'}_{\mu\nu} = -4\pi \mathrm{i} G \bar{\psi}^{A'}_{[\mu}\psi^A_{\nu]} , \tag{5.79}$$

where $S^{AA'}_{\mu\nu}$ denotes the torsion, and the last step follows from variation of the action with respect to the connection forms; see van Nieuwenhuizen (1981). The action (5.77) is invariant under the following infinitesimal local symmetry transformations:

1. *Supersymmetry (SUSY) transformations*:

$$\delta e_\mu^{AA'} = -\mathrm{i}\sqrt{8\pi G}(\epsilon^A \bar{\psi}_\mu^{A'} + \bar{\epsilon}^{A'} \psi_\mu^A), \tag{5.80}$$

$$\delta \psi_\mu^A = \frac{D_\mu \epsilon^A}{\sqrt{2\pi G}} \ , \quad \delta \bar{\psi}_\mu^{A'} = \frac{D_\mu \bar{\epsilon}^{A'}}{\sqrt{2\pi G}}, \tag{5.81}$$

where ϵ^A and $\bar{\epsilon}^{A'}$ denote anticommuting fields.

2. *Local Lorentz transformations*:

$$\delta e_\mu^{AA'} = N_B^A e_\mu^{BA'} + \bar{N}_{B'}^{A'} e^{AB'\mu}, \tag{5.82}$$

$$\delta \psi_\mu^A = N_B^A \psi_\mu^B \ , \quad \delta \bar{\psi}_\mu^{A'} = \bar{N}_{B'}^{A'} \bar{\psi}_\mu^{B'}, \tag{5.83}$$

with $N^{AB} = N^{(AB)}$.

3. *Local coordinate transformations*:

$$\delta e_\mu^{AA'} = \xi^\nu \partial_\nu e_\mu^{AA'} + e_\nu^{AA'} \partial_\mu \xi^\nu, \tag{5.84}$$

$$\delta \psi_\mu^A = \xi^\nu \partial_\nu \psi_\mu^A + \psi_\nu^A \partial_\mu \xi^\nu, \tag{5.85}$$

where ξ^ν are the parameters defining the (infinitesimal) coordinate transformation. The right-hand sides are just the Lie derivatives of these fields.

In analogy to Chapter 4 for GR, one can develop a Hamiltonian formalism for SUGRA. For this purpose, one splits $e_\mu^{AA'}$ into $e_0^{AA'}$ and $e_a^{AA'}$ to get the spatial metric

$$h_{ab} = -e_{AA'a} e_b^{AA'} = g_{ab}, \tag{5.86}$$

where $e_a^{AA'} = e_a^n \sigma_n^{AA'}$ in analogy to (5.74). The spinorial version of the normal vector n^μ reads

$$n^{AA'} = e_\mu^{AA'} n^\mu, \tag{5.87}$$

obeying

$$n_{AA'} e_a^{AA'} = 0 \ , \quad n_{AA'} n^{AA'} = 1. \tag{5.88}$$

Analogously to (4.40), one can expand the remaining components of the spinorial tetrad as

$$e_0^{AA'} = N n^{AA'} + N^a e_a^{AA'}, \tag{5.89}$$

with a lapse function N and shift vector N^a. The canonical formalism starts with the definition of the momenta. The momenta conjugate to N, N^a, ψ_0^A, and $\bar{\psi}_0^{A'}$ are all zero, since these variables are Lagrange multipliers. The momenta conjugate to the gravitino fields are[17]

$$\pi_A^a = \frac{\delta S}{\delta \dot{\psi}_a^A} = -\frac{1}{2}\epsilon^{abc} \bar{\psi}_b^{A'} e_{AA'c}, \tag{5.90}$$

$$\tilde{\pi}_{A'}^a = \frac{\delta S}{\delta \dot{\bar{\psi}}_a^{A'}} = \frac{1}{2}\epsilon^{abc} \psi_b^A e_{AA'c}. \tag{5.91}$$

Since the action is linear in $D\psi$ and $D\bar{\psi}$, the time derivatives $\dot{\psi}$ and $\dot{\bar{\psi}}$ do not occur on the right-hand sides. Therefore, these equations are in fact constraints. It turns

[17]The Grassmann-odd variables must be brought to the left before the functional differentiation is carried out. The momentum $\tilde{\pi}_{A'}^a$ is minus the Hermitian conjugate of π_A^a.

out that these constraints are of *second class*; that is, the Poisson brackets of the constraints do not close on the constraints again; cf. Section 3.1.2. As a consequence, one can eliminate the momenta π_A^a and $\tilde{\pi}_{A'}^a$ from the canonical action by using these constraints. Finally, the momentum conjugate to the spinorial tetrad can be found from

$$p_{AA'}^a = \frac{\delta S}{\delta \dot{e}_a^{AA'}}, \tag{5.92}$$

from which the ordinary spatial components follow via $p^{ab} = -e^{AA'a} p_{AA'}^b$. The symmetric part of p^{ab} can be expressed exactly as in (4.63) in terms of the second fundamental form K_{ab} on $t =$ constant,

$$p^{(ab)} = \frac{\sqrt{h}}{16\pi G} \left(K^{(ab)} - K h^{ab} \right). \tag{5.93}$$

However, due to the presence of torsion, K_{ab} now also possesses an antisymmetric part,

$$K_{[ab]} = S_{0ab} = n^\mu S_{\mu ab}. \tag{5.94}$$

If second-class constraints are present, one has to use *Dirac brackets* instead of Poisson brackets for the canonically conjugate variables (Dirac 1964, Sundermeyer 1982, Henneaux and Teitelboim 1992). Dirac brackets coincide with Poisson brackets on the constraint hypersurface but have the advantage that the variables of the original phase space can be used. They are denoted by $\{\ldots\}_*$. In the present case, one has

$$\begin{aligned}
\{e_a^{AA'}(x), e_b^{BB'}(y)\}_* &= 0, \\
\{e_a^{AA'}(x), p_{BB'}^b(y)\}_* &= \epsilon_B^A \epsilon_{B'}^{A'} \delta_a^b \delta(x,y), \\
\{p_{AA'}^a(x), p_{BB'}^b(y)\}_* &= \tfrac{1}{4}\, \epsilon^{bcd} \psi_{Bd} D_{AB'ec} \epsilon^{aef} \bar{\psi}_{A'f} \delta(x,y) + \text{h.c.}, \\
\{\psi_a^A(x), \psi_b^B(y)\}_* &= 0, \\
\{\psi_a^A(x), \bar{\psi}_b^{A'}(y)\}_* &= -D_{ab}^{AA'} \delta(x,y), \\
\{e_a^{AA'}(x), \psi_b^B(y)\}_* &= 0, \\
\{p_{AA'}^a(x), \psi_b^B(y)\}_* &= \tfrac{1}{2}\, \epsilon^{acd} \psi_{Ad} D_{A'bc}^B \delta(x,y),
\end{aligned} \tag{5.95}$$

where

$$D_{ab}^{AA'} = -\frac{2\mathrm{i}}{\sqrt{h}} e_b^{AB'} e_{BB'a} n^{BA'}. \tag{5.96}$$

In addition, one has the conjugate relations.

The invariance of the action under local Lorentz transformations yields the primary constraints

$$J_{AB} \approx 0\ , \quad \bar{J}_{A'B'} \approx 0, \tag{5.97}$$

where

$$J_{AB} = e_{(A}^{A'a} p_{B)A'a} + \psi_{(A}^a \pi_{B)a}. \tag{5.98}$$

In addition, one finds the secondary constraints

$$\mathcal{H}_\perp \approx 0\ , \quad \mathcal{H}_a \approx 0\ , \quad S_A \approx 0\ , \quad \bar{S}_{A'} \approx 0, \tag{5.99}$$

where S_A and $\bar{S}_{A'}$ denote the generators of SUSY transformations. It is also common to use the combination

$$\mathcal{H}_{AA'} = -n_{AA'}\mathcal{H}_\perp + e^a_{AA'}\mathcal{H}_a. \tag{5.100}$$

With the definition $\pi^{ab} = -p^{(ab)}/2$, one finds for the Hamiltonian constraint

$$\begin{aligned}\mathcal{H}_\perp = 16\pi G\ G_{abcd}\pi^{ab}\pi^{cd} - \frac{\sqrt{h}\,{}^{(3)}R}{16\pi G}\\ +\pi G\sqrt{h}n_{AA'}\bar{\psi}^{A'}_{[a}\psi^A_{b]}n^{BB'}\bar{\psi}^{[a}_{B'}\psi^{b]}_B\\ +\frac{1}{2}\epsilon^{abc}\bar{\psi}^{A'}_a n_{AA'}\mathcal{D}_b\psi^A_c + \text{h.c.}\end{aligned} \tag{5.101}$$

plus terms proportional to the Lorentz constraints. $\mathcal{D}_b$ denotes the three-dimensional version of (5.78),

$$\mathcal{D}_b\psi^A_a = \partial_b\psi^A_a + {}^{(3)}\omega^A_{Bb}\psi^B_a, \tag{5.102}$$

where ${}^{(3)}\omega^A_{Bb}$ are the spatial connection forms. For the explicit expressions for the other constraints, we refer to D'Eath (1984, 1996) and Moniz (2010); see below for the quantum versions of the Lorentz and the SUSY constraints.

The time evolution of a dynamical variable is given by

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \{A, H\}_*, \tag{5.103}$$

where the Hamiltonian is given by the expression

$$H = \int \mathrm{d}^3x\ \left(N\mathcal{H}_\perp + N^a\mathcal{H}_a + \psi^A_0 S_A + \bar{S}_{A'}\bar{\psi}^{A'}_0 - \omega_{AB0}J^{AB} - \bar{\omega}_{A'B'0}\bar{J}^{A'B'}\right), \tag{5.104}$$

from which the constraints follow after variation with respect to the Lagrange multipliers. This expression holds for the spatially compact case. In the asymptotically flat case, one again has terms at spatial infinity—the original ones (Section 4.2.4) plus supercharges at infinity arising from the global SUSY algebra (Section 2.3).

Of particular interest are the Dirac brackets among the SUSY generators, for which one finds

$$\{S_A(x), S_B(y)\}_* = 0\ , \quad \{\bar{S}_{A'}(x), \bar{S}_{B'}(y)\}_* = 0, \tag{5.105}$$

and

$$\{S_A(x), \bar{S}_{A'}(y)\}_* = 4\pi\mathrm{i}\, G\,\mathcal{H}_{AA'}(x)\delta(x, y) \tag{5.106}$$

plus terms proportional to the constraints (5.97). One recognizes from (5.106) that the constraints $\mathcal{H}_{AA'} \approx 0$ already follow from the validity of the remaining constraints. Since the SUSY constraints appear quadratically on the left and the Hamiltonian constraint linearly on the right, one can refer to $N = 1$ SUGRA as the 'square root of gravity' (Teitelboim 1977).

Quantization proceeds by turning Dirac brackets into commutators or anticommutators. Grassmann-even variables are quantized using commutators (omitting hats on operators),

$$\{E_1, E_2\}_* \longrightarrow -\frac{\mathrm{i}}{\hbar}[E_1, E_2], \tag{5.107}$$

while Grassmann-odd variables are quantized using anticommutators,

$$\{O_1, O_2\}_* \longrightarrow -\frac{\mathrm{i}}{\hbar}[O_1, O_2]_+. \tag{5.108}$$

Mixed variables are quantized via commutators,

$$\{O, E\}_* \longrightarrow -\frac{\mathrm{i}}{\hbar}[O, E]. \tag{5.109}$$

Proceeding as in Section 5.2.2, one can implement the Dirac brackets (5.95) via wave functionals

$$\Psi[e_a^{AA'}(x), \psi_a^A(x)], \tag{5.110}$$

which can depend either on $\psi_a^A(x)$ (as written here) or on $\bar{\psi}_a^{A'}$, but not both. This is because of the non-trivial anticommutator

$$[\psi_a^A(x), \bar{\psi}_b^{A'}(y)]_+ = -\mathrm{i}\hbar D_{ab}^{AA'}\delta(x, y), \tag{5.111}$$

which can be represented on the wave functional as[18]

$$\bar{\psi}_a^{A'}(x) = -\mathrm{i}\hbar D_{ba}^{AA'}\frac{\delta}{\delta\psi_b^A(x)}. \tag{5.112}$$

The momenta $p^a_{AA'}(x)$ are represented as

$$\begin{aligned} p^a_{AA'}(x) &= -\mathrm{i}\hbar\frac{\delta}{\delta e_a^{AA'}(x)} + \frac{1}{2}\epsilon^{abc}\psi_{Ab}(x)\bar{\psi}_{A'c}(x) \\ &= -\mathrm{i}\hbar\frac{\delta}{\delta e_a^{AA'}(x)} - \frac{\mathrm{i}\hbar}{2}\epsilon^{abc}\psi_{Ab}(x)D^B_{A'dc}\frac{\delta}{\delta\psi_d^B(x)}. \end{aligned} \tag{5.113}$$

The factor ordering has been chosen such that all derivatives are on the right. One can choose a formal Schrödinger-type inner product with respect to which $p^a_{AA'}$ is Hermitian and $\bar{\psi}_a^{A'}$, ψ_a^A are Hermitian adjoints (D'Eath 1996).[19]

The quantum constraints then read

$$J_{AB}\Psi = 0\ , \quad \bar{J}_{A'B'}\Psi = 0\ , \quad \mathcal{H}_{AA'}\Psi = 0\ , \quad S_A\Psi = 0\ , \quad \bar{S}_{A'}\Psi = 0. \tag{5.114}$$

The first two constraints express the invariance of the wave functional under Lorentz transformations, while the last two constraints express its invariance under SUSY transformations. From the quantum version of (5.106), it becomes clear that a solution

[18] One can also employ other representations; cf. the remarks in Moniz (2003).

[19] One can go via a functional Fourier transformation from Ψ to $\tilde{\Psi}[e_a^{AA'}(x), \bar{\psi}_a^{A'}(x)]$.

of the Lorentz and the SUSY constraints is *also* a solution to $\mathcal{H}_{AA'}\Psi = 0$, provided, of course, that there are no anomalies and the quantum algebra closes. The issue of anomalies is unsolved here as it is in the case of quantum GR.

The explicit form of the quantum Lorentz constraint operators reads

$$J_{AB} = -\frac{\mathrm{i}\hbar}{2}\left(e^{A'}_{Ba}\frac{\delta}{\delta e^{AA'}_a} + e^{A'}_{Aa}\frac{\delta}{\delta e^{BA'}_a} + \psi_{Ba}\frac{\delta}{\delta \psi^A_a} + \psi_{Aa}\frac{\delta}{\delta \psi^B_a}\right), \tag{5.115}$$

$$\bar{J}_{A'B'} = -\frac{\mathrm{i}\hbar}{2}\left(e^{A}_{B'a}\frac{\delta}{\delta e^{AA'}_a} + e^{A}_{A'a}\frac{\delta}{\delta e^{AB'}_a}\right), \tag{5.116}$$

while the quantum SUSY constraint operators read

$$\bar{S}_{A'} = \epsilon^{abc} e_{AA'a}\, {}^s\mathcal{D}_b \psi^A_c + 4\pi G\hbar \psi^A_a \frac{\delta}{\delta e^{AA'}_a}, \tag{5.117}$$

$$S_A = \mathrm{i}\hbar\, {}^s\mathcal{D}_a\left(\frac{\delta}{\delta \psi^A_a}\right) + 4\pi \mathrm{i} G\hbar \frac{\delta}{\delta e^{AA'}_a}\left(D^{BA'}_{ba}\frac{\delta}{\delta \psi^B_b}\right), \tag{5.118}$$

where ${}^s\mathcal{D}_a$ is the 'torsion-free derivative'. In the anomaly-free case, one has only to solve the constraints $J_{AB}\Psi = 0$, $S_A\Psi = 0$ (and their conjugates), since $\mathcal{H}_{AA'}\Psi = 0$ must hold automatically. This could lead to considerable simplification because (5.115)–(5.118) involve at most first-order derivatives.

Canonical quantum SUGRA can be applied, for example, in the context of quantum cosmology (Section 8.1). One can also study some general properties. One of them is the fact that pure bosonic states cannot exist (see e.g. Moniz 2010 for discussion and references). This can easily be shown. Considering a bosonic state $\Psi[e^{AA'}_a]$, one has $\delta\Psi/\delta\psi^A_a = 0$ and recognizes immediately that (5.118) is solved, that is, $S_A\Psi = 0$. Assuming that $\Psi[e^{AA'}_a]$ is Lorentz invariant, that is, that the Lorentz constraints are already fulfilled, it is clear that a state with $\bar{S}^{A'}\Psi = 0$ satisfies *all* constraints. However, such a state cannot exist. This can be seen as follows (Carroll *et al.* 1994). One multiplies $\bar{S}^{A'}\Psi = 0$ by $[\Psi]^{-1}$ and integrates over space with an arbitrary spinorial test function $\bar{\epsilon}^{A'}(x)$ to get

$$I := \int \mathrm{d}^3x\ \bar{\epsilon}^{A'}(x)\left(\epsilon^{abc} e_{AA'a}\, {}^s\mathcal{D}_b \psi^A_c + 4\pi G\hbar \psi^A_a \frac{\delta(\ln \Psi)}{\delta e^{AA'}_a}\right) = 0. \tag{5.119}$$

This must hold for all fields and all $\bar{\epsilon}^{A'}(x)$.

If one now replaces $\bar{\epsilon}^{A'}(x)$ by $\bar{\epsilon}^{A'}(x)\exp(-\phi(x))$ and $\psi^A_a(x)$ by $\psi^A_a(x)\exp(\phi(x))$, where $\phi(x)$ is some arbitrary function, the second term in (5.119) cancels out in the difference ΔI between the old and the new integral, and one is left with

$$\Delta I = -\int \mathrm{d}^3x\ \epsilon^{abc} e_{AA'a}\bar{\epsilon}^{A'}\psi^A_c \partial_b\phi = 0, \tag{5.120}$$

which is independent of the state Ψ. It is obvious that one can choose the fields, as well as $\bar{\epsilon}^{A'}(x)$ and $\phi(x)$, in such a way that the integral is non-vanishing, leading to a contradiction. Therefore, no physical bosonic states can exist, and a solution of the quantum constraints can be represented in the form

$$\Psi[e_a^{AA'}(x), \psi_a^A(x)] = \sum_{n=1}^{\infty} \Psi^{(n)}[e_a^{AA'}(x), \psi_a^A(x)], \tag{5.121}$$

where the expansion is into states with fermion number n. In fact, one can show by similar arguments that any solution of the quantum constraints must have an infinite fermion number. In this sense, pure quantum GR cannot arise as a limit from quantum SUGRA. An explicit solution of a peculiar type was found (without any regularization) by Csordás and Graham (1995).

5.4 The semiclassical approximation

5.4.1 Analogies from quantum mechanics

The semiclassical approximation to quantum geometrodynamics discussed here uses, in fact, a mixture of two different approximation schemes. The full system is divided into two parts with very different scales. One part is called the 'heavy part'—for it, the standard semiclassical (WKB) approximation is used. The other part is called the 'light part'—this is treated in a fully quantum way and follows the dynamics of the heavy part adiabatically. A mixed scheme of this kind is called a 'Born–Oppenheimer' type of approximation scheme. It is successfully applied in molecular physics, where the division is into the heavy nuclei (moving slowly) and the light electrons (following the motion of the nuclei adiabatically). Many molecular spectra can be explained in this way. In quantum gravity, the 'heavy' part is often taken to be the full gravitational field (motivated by the large value of the Planck mass), while the 'light' part is all non-gravitational degrees of freedom (see e.g. Kiefer 1994). This has the formal advantage that an expansion with respect to the Planck mass can be performed. It is, however, fully consistent to consider part of the gravitational degrees of freedom as fully quantum and therefore include them in the 'light' part (see e.g. Halliwell and Hawking 1985 and Vilenkin 1989). Physically, these degrees of freedom are gravitons and quantum density fluctuations. It depends on the actual situation one is interested in whether this 'light' gravitational part has to be taken into account or not.

For full quantum geometrodynamics, this semiclassical expansion exists only at a formal level. Therefore, it will be appropriate to discuss quantum-mechanical analogies in this subsection. Albeit formal, the expansion scheme is of the utmost conceptual importance. As we shall see, it enables one to recover the usual time as an approximate concept from 'timeless' quantum gravity (Section 5.4.2). This is the relevant approach for observers within the Universe (the 'intrinsic viewpont'). Moreover, the scheme allows one to advance to higher orders and calculate quantum-gravitational correction terms to the functional Schrödinger equation, which could have observational significance (Section 5.4.3).

Let us now consider in some detail a simple quantum-mechanical model (see e.g. Kiefer 1994, Bertoni *et al.* 1996, and Briggs and Rost 2001). The total system consists of the 'heavy part' described by the variable Q, while the 'light-part' variable is called q.[20] It is assumed that the full system is described by a stationary Schrödinger equation,

$$H\Psi(q,Q) = E\Psi(q,Q), \tag{5.122}$$

where the Hamilton operator is of the form

$$H = -\frac{\hbar^2}{2M}\frac{\partial^2}{\partial Q^2} + V(Q) + h(q,Q), \tag{5.123}$$

where $h(q,Q)$ contains the pure q-part and the interaction between q and Q. In the case of the Wheeler–DeWitt equation, the total energy is zero, $E = 0$. One now makes an expansion of the form

$$\Psi(q,Q) = \sum_n \chi_n(Q)\psi_n(q,Q) \tag{5.124}$$

and assumes that $\langle\psi_n|\psi_m\rangle = \delta_{nm}$ for each value of Q. The inner product used here is the ordinary $\mathcal{L}^2$ scalar product with respect to q only. It is only this part of the inner product that will be naturally available in quantum gravity, because it will correspond there to the standard non-gravitational degrees of freedom.

If we insert the ansatz (5.124) into the Schrödinger equation and project on $\psi_m(q,Q)$, we obtain

$$\sum_n \left\langle \psi_m \left| -\frac{\hbar^2}{2M}\frac{\partial^2}{\partial Q^2} \right| \psi_n \right\rangle \chi_n + V(Q)\chi_m \ + \sum_n \langle\psi_m|h|\psi_n\rangle\chi_n = E\chi_m, \tag{5.125}$$

where the Q-derivative acts on everything to the right. We now introduce the 'Born–Oppenheimer potentials'

$$\epsilon_{mn}(Q) := \langle\psi_m|h|\psi_n\rangle, \tag{5.126}$$

which for eigenstates of the 'light' variable, $h|\psi_n\rangle = \epsilon_n|\psi_n\rangle$, just read $\epsilon_{mn}(Q) = \epsilon_n(Q)\delta_{mn}$ (no summation). In molecular physics, this is often the case of interest. We shall, however, keep the formalism more general. We also introduce the quantity

$$A_{mn}(Q) := \mathrm{i}\hbar \left\langle \psi_m \left| \frac{\partial\psi_n}{\partial Q} \right.\right\rangle, \tag{5.127}$$

which serves as a connection in the corresponding momentum

$$\mathcal{P}_{mn} := \frac{\hbar}{\mathrm{i}}\left(\delta_{mn}\frac{\partial}{\partial Q} - \frac{\mathrm{i}}{\hbar}A_{mn}\right). \tag{5.128}$$

Making use of (5.126) and (5.128), one can write (5.125) in the form[21]

[20] For simplicity, the total system is considered to be two-dimensional. The extension to more dimensions is straightforward.

[21] Here, $\mathcal{P}^2_{mn} := \sum_k \mathcal{P}_{mk}\mathcal{P}_{kn}$.

$$\sum_n \left(\frac{\mathcal{P}^2_{mn}}{2M} + \epsilon_{mn}(Q) \right) \chi_n(Q) + V(Q)\chi_m(Q) = E\chi_m(Q). \tag{5.129}$$

The modification in the momentum $\mathcal{P}_{mn}$ and the 'Born–Oppenheimer potential' ϵ_{mn} express the 'back reaction' from the 'light' part onto the 'heavy part'. Inserting the ansatz (5.124) into the Schrödinger equation *without* projection on ψ_m, one gets

$$\sum_n \chi_n(Q) \left[h(q,Q) - \left(E - V(Q) + \frac{\hbar^2}{2M\chi_n} \frac{\partial^2 \chi_n}{\partial Q^2} \right) \right. \\ \left. - \frac{\hbar^2}{2M} \frac{\partial^2}{\partial Q^2} - \frac{\hbar^2}{M\chi_n} \frac{\partial \chi_n}{\partial Q} \frac{\partial}{\partial Q} \right] \psi_n(q,Q) = 0. \tag{5.130}$$

We emphasize that (5.129) and (5.130) are still *exact* equations, describing the coupling between the 'heavy' and the 'light' part.

Various approximations can now be performed. In the first step, one can assume that the 'heavy' part is approximately insensitive to changes in the 'light' part. This enables one to neglect the off-diagonal parts in (5.129), leading to

$$\left[\frac{1}{2M} \left(\frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial Q} - A_{nn}(Q) \right)^2 + V(Q) + E_n(Q) \right] \chi_n(Q) = E\chi_n(Q), \tag{5.131}$$

where $E_n(Q) := \epsilon_{nn}(Q)$. For real ψ_n, one can show that the connection vanishes, $A_{nn} = 0$. Otherwise, it leads to a geometric phase ('Berry phase'); cf. Berry (1984). We shall neglect the connection in the following.[22]

In the second step, one can perform a standard semiclassical (WKB) approximation for the heavy part through the ansatz

$$\chi_n(Q) = C_n(Q) \mathrm{e}^{\mathrm{i}MS_n(Q)/\hbar}. \tag{5.132}$$

This is inserted into (5.131). For the Q-derivative, one gets

$$\frac{\partial^2 \chi_n}{\partial Q^2} = \frac{\partial^2 C_n}{\partial Q^2} \frac{\chi_n}{C_n} + \frac{2\mathrm{i}M}{\hbar} \frac{\partial C_n}{\partial Q} \frac{\partial S_n}{\partial Q} \frac{\chi_n}{C_n} \\ - \left(\frac{M}{\hbar} \right)^2 \left(\frac{\partial S_n}{\partial Q} \right)^2 \chi_n + \frac{\mathrm{i}M}{\hbar} \frac{\partial^2 S_n}{\partial Q^2} \chi_n. \tag{5.133}$$

Assuming M to be large corresponds to neglecting derivatives of C_n and second derivatives of S_n (the usual assumptions for the WKB approximation). One then has

$$\frac{\partial^2 \chi_n}{\partial Q^2} \approx - \left(\frac{M}{\hbar} \right)^2 \left(\frac{\partial S_n}{\partial Q} \right)^2 \chi_n. \tag{5.134}$$

The classical momentum is then given by

[22] An intriguing idea would be to derive the connection in gauge theories along these lines.

$$P_n = M\frac{\partial S_n}{\partial Q} \approx \frac{\hbar}{\mathrm{i}\chi_n}\frac{\partial \chi_n}{\partial Q}, \tag{5.135}$$

and (5.131) becomes the Hamilton–Jacobi equation,

$$H_{\mathrm{cl}} := \frac{P_n^2}{2M} + V(Q) + E_n(Q) = E. \tag{5.136}$$

Since $E_n(Q) = \langle \psi_n | h | \psi_n \rangle$, this corresponds, in the gravitational context, to the semiclassical Einstein equations discussed in Section 1.2, where the expectation value of the energy–momentum tensor appears.

One can now introduce a time coordinate t via the Hamilton equations of motion,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}P_n &= -\frac{\partial}{\partial Q}H_{\mathrm{cl}} = -\frac{\partial}{\partial Q}(V(Q) + E_n(Q)), \\ \frac{\mathrm{d}}{\mathrm{d}t}Q &= \frac{\partial}{\partial P_n}H_{\mathrm{cl}} = \frac{P_n}{M}. \end{aligned} \tag{5.137}$$

In fact, the very definition of t depends on n, and we therefore call it t_n in the following. Since it arises from the WKB approximation (5.132), it is called the *WKB time* (Zeh 1988). The last term in (5.130) can then be written as

$$-\frac{\hbar^2}{M\chi_n}\frac{\partial \chi_n}{\partial Q}\frac{\partial \psi_n}{\partial Q} \approx -\mathrm{i}\hbar\frac{\partial S_n}{\partial Q}\frac{\partial \psi_n}{\partial Q} =: -\mathrm{i}\hbar\frac{\partial \psi_n}{\partial t_n}. \tag{5.138}$$

This means that ψ_n is evaluated along a particular classical trajectory of the 'heavy' variable, $\psi_n(Q(t_n), q) \equiv \psi_n(t_n, q)$. Assuming slow variation of ψ_n with respect to Q, one can neglect the term proportional to $\partial^2\psi_n/\partial Q^2$ in (5.130). Using also (5.136), one is then left with

$$\sum_n \chi_n \left[h(q, t_n) - E_n(t_n) - \mathrm{i}\hbar\frac{\partial}{\partial t_n} \right] \psi_n(t_n, q) = 0. \tag{5.139}$$

This equation still describes a coupling between a 'heavy' and a 'light' part.

In the third and last step, one can assume that instead of the whole sum (5.124) only one component is available; that is, one has—up to an (adiabatic) dependence of ψ on Q—a factorizing state,

$$\chi(Q)\psi(q, Q).$$

This lack of entanglement can of course only arise in certain situations and must be dynamically justified (through decoherence; cf. Chapter 10). If it happens, and after absorbing $E_n(t)$ into a redefinition of ψ (yielding only a phase), one gets from (5.139)

$$\mathrm{i}\hbar\frac{\partial \psi}{\partial t} = h\psi, \tag{5.140}$$

that is, just the Schrödinger equation. The 'heavy' system acts as a 'clock' and defines the time with respect to which the 'light' system evolves. Therefore, a time-dependent Schrödinger equation has arisen for one of the subsystems, although the full Schrödinger equation is of a stationary form; cf. Mott (1931).

Considering the terms with order M in (5.133), one finds an equation for the C_n,

$$2\frac{\partial C_n}{\partial Q}\frac{\partial S_n}{\partial Q}+\frac{\partial^2 S_n}{\partial Q^2}C_n=0, \tag{5.141}$$

or, in the case of one component only,

$$2\frac{\partial S}{\partial Q}\frac{\partial C}{\partial Q}+\frac{\partial^2 S}{\partial Q^2}C=2\frac{\partial C}{\partial t}+\frac{\partial^2 S}{\partial Q^2}C=0. \tag{5.142}$$

This can be written in the form of a continuity equation,

$$\frac{\partial}{\partial Q}\left(C^2\frac{\partial S}{\partial Q}\right)=0. \tag{5.143}$$

A systematic derivation makes use of an M-expansion,

$$\Psi := \exp\left(\mathrm{i}\mathcal{S}/\hbar\right)\ , \quad \mathcal{S}=MS_0+S_1+M^{-1}S_2+\cdots . \tag{5.144}$$

The Hamilton–Jacobi equation (5.136) then appears at order M, and both the Schrödinger equation (5.140) and the prefactor equation (5.143) appear at order M^0. The next order, M^{-1}, then yields corrections to the Schrödinger equation (discussed in Section 5.4.3 for the quantum-gravitational case).

Another example, which is closer to the situation for the Wheeler–DeWitt equation (because of the indefinite kinetic term), is the non-relativistic expansion of the Klein–Gordon equation; cf. Kiefer and Singh (1991). In the absence of an external potential, the latter reads

$$\left(\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2}-\hbar^2\Delta+m^2c^2\right)\varphi(\mathbf{x},t)=0. \tag{5.145}$$

One writes

$$\varphi(\mathbf{x},t) := \exp\left(\mathrm{i}S(\mathbf{x},t)/\hbar\right) \tag{5.146}$$

and expands the exponent in powers of c,

$$S=c^2S_0+S_1+c^{-2}S_2+\cdots \tag{5.147}$$

This ansatz is inserted into (5.145), which is then solved at consecutive orders of c^2. The highest order, c^2, yields solutions $S_0=\pm mt$, from which we choose $S_0=-mt$. This choice corresponds to

$$\varphi=\exp\left(-\mathrm{i}mc^2t/\hbar\right),$$

which is the 'positive-energy solution'. The approximation works as long as 'negative-energy solutions' can be consistently neglected, that is, as long as field-theoretic effects such as particle creation do not play any role. Writing $\psi := \exp(\mathrm{i}S_1/\hbar)$, one gets at order c^0 the Schrödinger equation,

$$\mathrm{i}\hbar\dot{\psi} = -\frac{\hbar^2}{2m}\Delta\psi. \tag{5.148}$$

The Schrödinger equation has thus been derived as a non-relativistic approximation to the Klein–Gordon equation. Writing $\Psi := \psi \exp(\mathrm{i}S_2/\hbar c^2)$, one arrives at order c^{-2} at the equation

$$\mathrm{i}\hbar\dot{\Psi} = -\frac{\hbar^2}{2m}\Delta\Psi - \frac{\hbar^4}{8m^3c^2}\Delta\Delta\Psi. \tag{5.149}$$

The last term is the first relativistic correction term and can be used to derive testable predictions, for example for spectra of pionic atoms (for ordinary atoms, an expansion of the Dirac equation must be employed). A more general case is the Klein–Gordon equation coupled to gravity and the electromagnetic field. This leads to additional relativistic correction terms (Lämmerzahl 1995).

A major difference between the Klein–Gordon example and the first example is the indefinite structure of the kinetic term (d'Alembertian instead of Laplacian). Therefore, on the full level, the conserved inner product is the Klein–Gordon one (cf. Section 5.2.2). At order c^0 of the approximation, one obtains from this inner product the standard Schrödinger inner product as an approximation. The next order yields corrections to the Schrödinger inner product proportional to c^{-2}. Does this mean that unitarity is violated at this order? Not necessarily. In the case of the Klein–Gordon equation in external gravitational and electromagnetic fields, one can make a (t-dependent!) redefinition of the wave functions and Hamiltonian to arrive at a conserved Schrödinger inner product with respect to which the Hamiltonian is Hermitian (Lämmerzahl 1995).

5.4.2 Derivation of the Schrödinger equation

As in the discussion of the examples in the last subsection, one can perform a semiclassical ('Born–Oppenheimer') approximation for the Wheeler–DeWitt equation and the momentum constraints. In this way, one can recover approximately the limit of ordinary quantum field theory in an external gravitational background. This is done in the Schrödinger picture, so this limit emerges through the *functional* Schrödinger equation, not the quantum-mechanical Schrödinger equation as in the last subsection. In the following, we shall mainly follow, with elaborations, the presentation in Barvinsky and Kiefer (1998); see also Kiefer (1994) and references therein.

The starting point is the Wheeler–DeWitt equation (5.18) and the momentum constraint (5.19). Taking into account non-gravitational degrees of freedom, these equations can be written in the following form:

$$\left\{-\frac{1}{2m_{\mathrm{P}}^2}G_{abcd}\frac{\delta^2}{\delta h_{ab}\delta h_{cd}} - 2m_{\mathrm{P}}^2\sqrt{h}\,{}^{(3)}\!R + \hat{\mathcal{H}}_{\perp}^{\mathrm{m}}\right\}|\Psi[h_{ab}]\rangle = 0, \tag{5.150}$$

$$\left\{-\frac{2}{\mathrm{i}}h_{ab}D_c\frac{\delta}{\delta h_{bc}} + \hat{\mathcal{H}}_a^{\mathrm{m}}\right\}|\Psi[h_{ab}]\rangle = 0. \tag{5.151}$$

Here, we define $m_{\rm P}^2 := (32\pi G)^{-1}$, differently from the rest of the book, and choose $\hbar = 1$ and $\Lambda = 0$; $\hat{\mathcal{H}}_\perp^{\rm m}$ and $\hat{\mathcal{H}}_a^{\rm m}$ denote the contributions from non-gravitational fields. To be concrete, we think about the presence of a scalar field. The notation $|\Psi[h_{ab}]\rangle$ means that Ψ is a wave functional with respect to the three-metric h_{ab} and a state in the standard Hilbert space referring to the scalar field (using bra and ket notation).

The situation is now formally similar to the example discussed in the previous subsection. One of the main differences is the presence of the momentum constraints (5.151), which have no analogue in the quantum-mechanical example. Comparing (5.150) with (5.123), one notes the following formal correspondence between the terms:

$$\begin{aligned} -\frac{1}{2M}\frac{\partial^2}{\partial Q^2} &\leftrightarrow -\frac{1}{2m_{\rm P}^2}G_{abcd}\frac{\delta^2}{\delta h_{ab}\delta h_{cd}}, \\ V(Q) &\leftrightarrow -2m_{\rm P}^2\sqrt{h}\,{}^{(3)}\!R, \\ h(q,Q) &\leftrightarrow \hat{\mathcal{H}}_\perp^{\rm m}, \\ \Psi(q,Q) &\leftrightarrow |\Psi[h_{ab}]\rangle. \end{aligned} \tag{5.152}$$

The same steps as in the quantum-mechanical example can now be performed. We have already assumed at this stage that we have one component instead of a sum like (5.124) and that this component is written in the form

$$|\Psi[h_{ab}]\rangle = C[h_{ab}]{\rm e}^{{\rm i}m_{\rm P}^2 S[h_{ab}]}|\psi[h_{ab}]\rangle. \tag{5.153}$$

At the highest order of a WKB approximation for the gravitational part, $S[h_{ab}]$ obeys a Hamilton–Jacobi equation similar to (5.136). In addition, it obeys a Hamilton–Jacobi version of the momentum constraints. Therefore,

$$\frac{m_{\rm P}^2}{2}G_{abcd}\frac{\delta S}{\delta h_{ab}}\frac{\delta S}{\delta h_{cd}} - 2m_{\rm P}^2\sqrt{h}\,{}^{(3)}\!R + \langle\psi|\hat{\mathcal{H}}_\perp^{\rm m}|\psi\rangle = 0, \tag{5.154}$$

$$-2m_{\rm P}^2 h_{ab}D_c\frac{\delta S}{\delta h_{bc}} + \langle\psi|\hat{\mathcal{H}}_a^{\rm m}|\psi\rangle = 0. \tag{5.155}$$

These equations correspond—in the usual four-dimensional notation—to the semiclassical Einstein equations (1.37). Note that the 'back reaction' terms in these equations are formally suppressed by a factor $m_{\rm P}^{-2}$ compared to the remaining terms. For this reason, they are often neglected at this order of approximation and only considered at the next order (Kiefer 1994).

The analogue of (5.139) consists of the equations

$$\left(\hat{\mathcal{H}}_\perp^{\rm m} - \langle\psi|\hat{\mathcal{H}}_\perp^{\rm m}|\psi\rangle - {\rm i}G_{abcd}\frac{\delta S}{\delta h_{ab}}\frac{\delta}{\delta h_{cd}}\right)|\psi[h_{ab}]\rangle = 0, \tag{5.156}$$

$$\left(\hat{\mathcal{H}}_a^{\rm m} - \langle\psi|\hat{\mathcal{H}}_a^{\rm m}|\psi\rangle - \frac{2}{\rm i}h_{ab}D_c\frac{\delta}{\delta h_{bc}}\right)|\psi[h_{ab}]\rangle = 0. \tag{5.157}$$

One now evaluates $|\psi[h_{ab}]\rangle$ along a solution of the classical Einstein equations, $h_{ab}(\mathbf{x},t)$, corresponding to a solution, $S[h_{ab}]$, of the Hamilton–Jacobi equations (5.154) and (5.155),

$$|\psi(t)\rangle = |\psi[h_{ab}(\mathbf{x},t)]\rangle. \tag{5.158}$$

After a certain *choice* of the lapse and shift functions N and N^a has been made, this solution is obtained from

$$\dot{h}_{ab} = NG_{abcd}\frac{\delta S}{\delta h_{cd}} + 2D_{(a}N_{b)}. \tag{5.159}$$

To get an evolutionary equation for the quantum state (5.158), one defines

$$\frac{\partial}{\partial t}|\psi(t)\rangle = \int \mathrm{d}^3x\, \dot{h}_{ab}(\mathbf{x},t)\,\frac{\delta}{\delta h_{ab}(\mathbf{x})}|\psi[h_{ab}]\rangle. \tag{5.160}$$

This then leads to the functional Schrödinger equation for quantized matter fields in the chosen external classical gravitational field,

$$\begin{aligned}\mathrm{i}\frac{\partial}{\partial t}|\psi(t)\rangle &= \hat{H}^{\mathrm{m}}|\psi(t)\rangle,\\ \hat{H}^{\mathrm{m}} &:= \int \mathrm{d}^3x\left\{N(\mathbf{x})\hat{\mathcal{H}}_\perp^{\mathrm{m}}(\mathbf{x}) + N^a(\mathbf{x})\hat{\mathcal{H}}_a^{\mathrm{m}}(\mathbf{x})\right\}.\end{aligned} \tag{5.161}$$

Here, $\hat{H}^{\mathrm{m}}$ is the matter-field Hamiltonian in the Schrödinger picture, parametrically depending on (generally non-static) metric coefficients of the curved space–time background. This equation is the analogue of (5.140) in the quantum-mechanical example. (The back-reaction terms have again been absorbed into the phase of $|\psi(t)\rangle$.) The standard concept of time in quantum theory thus emerges here only in a semiclassical approximation—the Wheeler–DeWitt equation itself is 'timeless'.[23]

Such a derivation of quantum field theory from the Wheeler–DeWitt equations dates back, on the level of cosmological models, to DeWitt (1967*a*). It was later performed by Lapchinsky and Rubakov (1979) for generic gravitational systems and discussed in various contexts by Banks (1985), Halliwell and Hawking (1985), Hartle (1987), Kiefer (1987), Barvinsky (1989), Brout and Venturi (1989), Singh and Padmanabhan (1989), Parentani (2000), and others; see also Anderson (2007*b*,*c*) for a general discussion. Although performed at a formal level only, this derivation yields an important bridge connecting the full theory of quantum gravity with the limit of quantum field theory in an external space–time; it lies behind many cosmological applications.

This 'Born–Oppenheimer type of approach' is also well suited for the calculation of quantum-gravitational correction terms to the Schrödinger equation (5.161). This will be discussed in the next subsection. As a preparation it is, however, most appropriate to introduce again a condensed, so-called 'DeWitt', notation; cf. Section 2.2. We introduce the notation

$$q^i = h_{ab}(\mathbf{x})\ , \quad p_i = p^{ab}(\mathbf{x}), \tag{5.162}$$

in which the condensed index $i = (ab, \mathbf{x})$ includes both discrete tensor indices and three-dimensional spatial coordinates $\mathbf{x}$. In this way, the situation is formally the

[23] An attempt has been made to extrapolate the standard interpretational framework of quantum theory into the 'timeless realm' by the use of 'evolving constants' in the Heisenberg picture by Rovelli (1991*b*).

same as for a finite-dimensional model. A similar notation can be introduced for the constraints,

$$H_\mu^{\mathrm{g}}(q,p) = (\mathcal{H}_\perp^{\mathrm{g}}(\mathbf{x}), \mathcal{H}_a^{\mathrm{g}}(\mathbf{x}))\ , \quad H_\mu^{\mathrm{m}}(q,\varphi,p_\varphi) = (\mathcal{H}_\perp^{\mathrm{m}}(\mathbf{x}), \mathcal{H}_a^{\mathrm{m}}(\mathbf{x})). \tag{5.163}$$

The index μ enumerates the super-Hamiltonian and supermomenta of the theory as well as their spatial coordinates, $\mu \to (\mu, \mathbf{x})$. In this notation, the functional dependence on phase-space variables is represented in the form of functions of (q^i, p_i), and the contraction of condensed indices includes integration over $\mathbf{x}$ along with discrete summation. In the condensed notation, the gravitational part of the canonical action (4.68) acquires the simple form

$$S^{\mathrm{g}}[q,p,N] = \int \mathrm{d}t\ \left(p_i \dot{q}^i - N^\mu H_\mu^{\mathrm{g}}(q,p)\right) =: \int \mathrm{d}t\ L^{\mathrm{g}}, \tag{5.164}$$

with the super-Hamiltonian and supermomenta given by expressions which are quadratic and linear, respectively, in the momenta:

$$H_\perp^{\mathrm{g}} = \frac{1}{2m_{\mathrm{P}}^2} G_\perp^{ik} p_i p_k + V_\perp\ , \quad H_a^{\mathrm{g}} = D_a^i p_i, \tag{5.165}$$

with $V_\perp = -2m_{\mathrm{P}}^2 \sqrt{h}\,{}^{(3)}R$. Here the indices $\perp \to (\perp, \mathbf{x})$ and $a \to (a, \mathbf{x})$ are also condensed, $G_\perp^{ik}$ is the ultralocal three-point object containing the matrix of the DeWitt metric, and D_a^i is the generator of the spatial diffeomorphisms (see below). The objects $G_\perp^{ik}$ and D_a^i have the form of the following delta-function-type kernels (Barvinsky 1993*b*):

$$\begin{aligned} G_\perp^{ik} &= G_{abcd}\ \delta(\mathbf{x}_i, \mathbf{x}_k)\delta(\mathbf{x}_\perp, \mathbf{x}_k), \\ &\qquad i = (ab, \mathbf{x}_i),\ \ k = (cd, \mathbf{x}_k),\ \ \perp = (\perp, \mathbf{x}_\perp), \end{aligned} \tag{5.166}$$

$$D_a^i = -2h_{a(b}D_{c)}\delta(\mathbf{x}_a, \mathbf{x}_i), \quad i = (bc, \mathbf{x}_i),\ \ a = (a, \mathbf{x}_a). \tag{5.167}$$

Note that the object $G_\perp^{ik}$ itself is not yet the DeWitt metric, because it contains two delta functions. Only the functional contraction of $G_\perp^{ik}$ with the constant lapse function $N = 1$ converts this object into the distinguished ultralocal metric on the functional space of three-metric coefficients,

$$G^{ik} = G_\perp^{ik} N\big|_{N=1} \equiv \int \mathrm{d}^3\mathbf{x}_\perp\ G_\perp^{ik} = G_{abcd}\ \delta(\mathbf{x}_i, \mathbf{x}_k). \tag{5.168}$$

The Poisson-bracket algebra for the gravitational constraints in condensed notation then reads as in (5.51).

Note that the transformations of the q^i part of phase space (cf. Section 5.3.4)

$$\delta q^i = D_\mu^i \epsilon^\mu\ , \quad D_\mu^i := \frac{\partial H_\mu^{\mathrm{g}}}{\partial p_i} \tag{5.169}$$

have as generators the vectors D_μ^i, which are momentum-independent for space-like diffeomorphisms $\mu = a$ (and, therefore, coincide with the coefficients of the momenta in the supermomentum constraints (5.167)), but involve momenta for normal deformations, $D^i_\perp = G^{ik}_\perp p_k / m_\mathrm{P}^2$.

With these condensed notations, one can formulate the operator realization of the gravitational constraints $H^\mathrm{g}_\mu(q,p) \to \hat{H}^\mathrm{g}_\mu$, closing the commutator version of the Poisson-bracket algebra (5.51),

$$\left[\hat{H}^\mathrm{g}_\mu, \hat{H}^\mathrm{g}_\nu\right] = \mathrm{i}\hat{U}^\lambda_{\mu\nu}\hat{H}^\mathrm{g}_\lambda. \tag{5.170}$$

As shown in Barvinsky (1993*b*), the validity of (5.170) follows from the classical gravitational constraints (5.165) if the following two manipulations are done. First, the momenta p_k are replaced by the functional covariant derivatives $\mathcal{D}_k/\mathrm{i}$, where covariance refers to the Riemann connection based on the DeWitt metric (5.168). Second, a purely imaginary part (which is anti-Hermitian with respect to the $\mathcal{L}^2$ inner product) is added: the functional trace of structure functions, $\mathrm{i}U^\nu_{\mu\nu}/2$. With this definition of covariant derivatives, it is understood that the space of three-metrics q is regarded as a functional differentiable manifold, and that the quantum states $|\Psi(q)\rangle$ are scalar densities of weight 1/2. Thus, the operator realization for the full constraints including the matter parts has the form

$$\hat{H}_\perp = -\frac{1}{2m_\mathrm{P}^2}G^{ik}_\perp \mathcal{D}_i\mathcal{D}_k + V_\perp + \frac{\mathrm{i}}{2}U^\nu_{\perp\nu} + \hat{H}^\mathrm{m}_\perp, \tag{5.171}$$

$$\hat{H}_a = \frac{1}{\mathrm{i}}\nabla^i_a\mathcal{D}_i + \frac{\mathrm{i}}{2}U^\nu_{a\nu} + \hat{H}^\mathrm{m}_a. \tag{5.172}$$

The imaginary parts of these operators are either formally divergent (being the coincidence limits of delta-function-type kernels) or formally zero (as in (5.171) because of vanishing structure-function components). We shall, however, keep them in a general form, expecting that a rigorous operator regularization will exist that can consistently handle these infinities as well as the corresponding quantum anomalies (see Section 5.3.5).

The highest order of the semiclassical approximation leads to a wave functional of the form (5.153). It is convenient here to consider a two-point object ('propagator') instead of a wave functional. The reason is that one can then easily translate the results into a language using Feynman diagrams; this will be done in the next subsection. We shall, therefore, consider a two-point solution $\boldsymbol{K}(q,q')$ of the Wheeler–DeWitt equation, but want to mention that most of the following expressions can easily be rewritten in terms of wave functionals.

The solution that is of the semiclassical form corresponding to (5.153) reads (Barvinsky and Krykhtin 1993)

$$\boldsymbol{K}(q,q') = \boldsymbol{P}(q,q')\,\mathrm{e}^{\mathrm{i}m_\mathrm{P}^2\boldsymbol{S}(q,q')}. \tag{5.173}$$

Here, $\boldsymbol{S}(q,q')$ is a particular solution of the Hamilton–Jacobi equations with respect to both arguments—the classical action calculated at the extremal of equations of motion joining the points q and q',

$$H^\mathrm{g}_\mu(q, \partial\boldsymbol{S}/\partial q) = 0. \tag{5.174}$$

The one-loop order (the $O(m_{\rm P}^0)$ part of the $1/m_{\rm P}^2$-expansion) of the pre-exponential factor $\boldsymbol{P}(q,q')$ here satisfies a set of quasi-continuity equations which follow from the Wheeler–DeWitt equations at one loop and are analogous to (5.143) in the quantum-mechanical example,

$$\mathcal{D}_i(D^i_\mu \boldsymbol{P}^2) = U^\lambda_{\mu\lambda}\boldsymbol{P}^2, \tag{5.175}$$

$$D^i_\mu := \left.\frac{\partial H^{\rm g}_\mu}{\partial p_i}\right|_{p=\partial \boldsymbol{S}/\partial q}, \tag{5.176}$$

with the generators D^i_μ *here* evaluated at the Hamilton–Jacobi values of the canonical momenta. The solution of this equation turns out to be a particular generalization of the Pauli–van Vleck–Morette formula—the determinant calculated on the subspace of non-degeneracy for the matrix

$$\boldsymbol{S}_{ik'} = \frac{\partial^2 \boldsymbol{S}(q,q')}{\partial q^i\,\partial q^{k'}}. \tag{5.177}$$

This matrix has the generators D^i_μ as zero-eigenvalue eigenvectors (Barvinsky and Krykhtin 1993). An invariant algorithm for calculating this determinant is equivalent to the Faddeev–Popov gauge-fixing procedure; cf. Section 2.2.3. It consists of introducing a 'gauge-breaking' term into the matrix (5.177),

$$\boldsymbol{F}_{ik'} = \boldsymbol{S}_{ik'} + \phi^\mu_i c_{\mu\nu}\phi^\nu_{k'}, \tag{5.178}$$

formed with the aid of the gauge-fixing matrix $c_{\mu\nu}$ and two sets of arbitrary covectors (of 'gauge conditions') ϕ^μ_i and $\phi^\nu_{k'}$ at the points q and q', respectively. They satisfy invertibility conditions for 'Faddeev–Popov operators' at these two points,

$$J^\mu_\nu = \phi^\mu_i D^i_\nu\ ,\ \ J := \det J^\mu_\nu \neq 0\ ,\ \ J'^\mu_\nu = \phi^\mu_{i'} D^{i'}_\nu\ ,\ \ J' := \det J'^\mu_\nu \neq 0. \tag{5.179}$$

In terms of these objects, the pre-exponential factor solving the continuity equations (5.175) is given by

$$\boldsymbol{P} = \left[\frac{\det \boldsymbol{F}_{ik'}}{JJ'\det c_{\mu\nu}}\right]^{1/2}, \tag{5.180}$$

which is independent of the gauge fixing. One thus has found a closed expression for the one-loop pre-exponential factor. This finishes the discussion of the Born–Oppenheimer scheme at the highest level of approximation.

5.4.3 Quantum-gravitational correction terms

We shall now proceed to perform the semiclassical expansion for solutions to the Wheeler–DeWitt equations. Since we are interested in giving an interpretation in terms of Feynman diagrams, we shall consider not wave functionals but—as in the last part of the last subsection—two-point solutions ('propagators'). Due to the absence of an external time parameter in the full theory, such two-point functions play more the role of energy Green functions than ordinary propagators; see Section 5.3.4. In the semiclassical limit, however, a background time parameter is available, with respect to which Feynman 'propagators' can be formulated.

Let us, therefore, look for a two-point solution of the Wheeler–DeWitt equations (i.e. the Wheeler–DeWitt equation and momentum constraints) in the form of the ansatz

$$\hat{\boldsymbol{K}}(q_+, q_-) = \boldsymbol{P}(q_+, q_-)\mathrm{e}^{\mathrm{i}m_{\mathrm{P}}^2 \boldsymbol{S}(q_+,q_-)}\hat{\boldsymbol{U}}(q_+, q_-), \tag{5.181}$$

where, as above, we denote by a hat the operators acting in the Hilbert space of matter fields. Here, $\boldsymbol{S}(q_+, q_-)$ satisfies (5.174), and $\boldsymbol{P}(q_+, q_-)$ is the pre-exponential factor (5.180). Substituting this ansatz into the system of the Wheeler–DeWitt equations and taking into account the Hamilton–Jacobi equations and the continuity equations for $\boldsymbol{P}(q_+, q_-)$, we get for the 'evolution' operator $\hat{\boldsymbol{U}}(q_+, q_-)$ the equations

$$\mathrm{i}D_\perp^k \mathcal{D}_k \hat{\boldsymbol{U}} = \hat{H}_\perp^{\mathrm{m}} \hat{\boldsymbol{U}} - \frac{1}{2m_{\mathrm{P}}^2}\boldsymbol{P}^{-1} G_\perp^{mn} \mathcal{D}_m \mathcal{D}_n (\boldsymbol{P}\hat{\boldsymbol{U}}), \tag{5.182}$$

$$\mathrm{i}D_a^k \mathcal{D}_k \hat{\boldsymbol{U}} = \hat{H}_a^{\mathrm{m}} \hat{\boldsymbol{U}}, \tag{5.183}$$

where all the derivatives are understood as acting on the argument q_+. Evaluating this operator at the classical extremal $q_+ \to q(t)$,

$$\hat{\boldsymbol{U}}(t) = \hat{\boldsymbol{U}}(q(t), q_-), \tag{5.184}$$

where $q(t)$ satisfies the canonical equations of motion corresponding to $\boldsymbol{S}(q_+, q_-)$,

$$\dot{q}^i = N^\mu \nabla_\mu^i, \tag{5.185}$$

one easily obtains the quasi-evolutionary equation

$$\mathrm{i}\frac{\partial}{\partial t}\hat{\boldsymbol{U}}(t) = \hat{H}^{\mathrm{eff}} \hat{\boldsymbol{U}}(t) \tag{5.186}$$

with the *effective* matter Hamiltonian

$$\hat{H}^{\mathrm{eff}} = \hat{H}^{\mathrm{m}} - \frac{1}{2m_{\mathrm{P}}^2} N G_\perp^{mn} \mathcal{D}_m \mathcal{D}_n [\,\boldsymbol{P}\hat{\boldsymbol{U}}\,] \boldsymbol{P}^{-1} \hat{\boldsymbol{U}}^{-1}. \tag{5.187}$$

(Recall that we are using a condensed notation and that this equation is, in fact, an *integral* equation.) The first term on the right-hand side is the Hamiltonian of matter fields on the gravitational background of the (q, N) variables,

$$\hat{H}^{\mathrm{m}} = N^\mu \hat{H}_\mu^{\mathrm{m}}. \tag{5.188}$$

The second term involves the operator $\hat{\boldsymbol{U}}$ itself in a non-linear way and contributes only at order m_{P}^{-2} of the expansion. Thus, (5.186) is not a true linear Schrödinger equation, but semiclassically it can be solved by iteration starting from the lowest-order approximation

$$\hat{\boldsymbol{U}}_0 = T \exp\left[-\mathrm{i}\int_{t_-}^{t_+} \mathrm{d}t\, \hat{H}^{\mathrm{m}}\right], \tag{5.189}$$

$$\hat{H}_0^{\mathrm{eff}} = \hat{H}^{\mathrm{m}}. \tag{5.190}$$

Here, T denotes the Dyson chronological ordering of the usual unitary evolution operator acting in the Hilbert space of matter fields $(\hat{\varphi}, \hat{p}_\varphi)$. The Hamiltonian $\hat{H}^{\mathrm{m}} = H^{\mathrm{m}}(\hat{\varphi}, \hat{p}_\varphi, q(t), N(t))$ is an operator in the Schrödinger picture of these fields $(\hat{\varphi}, \hat{p}_\varphi)$, parametrically depending on the gravitational background variables $(q(t), N(t))$, that is, evaluated along a *particular* trajectory ('space–time') in configuration space.

The Dyson T-exponent obviously explains the origin of the standard Feynman diagrammatic technique in the matter field sector of the theory; this technique thus arises only in the course of the semiclassical expansion of (5.189). We shall now show that the gravitational part of this diagrammatic technique, involving graviton loops, arise naturally as a result of solving (5.186)–(5.187) iteratively in powers of $1/m_{\mathrm{P}}^2$.

The effective Hamiltonian in the first-order approximation of such an iterative technique can be obtained by substituting (5.189) into (5.187) to yield

$$\begin{aligned}\hat{H}_1^{\mathrm{eff}}(t_+) = \hat{H}^{\mathrm{m}} &- \frac{1}{2m_{\mathrm{P}}^2}\mathcal{G}^{mn}(\mathcal{D}_m\mathcal{D}_n\hat{\boldsymbol{U}}_0)\hat{\boldsymbol{U}}_0{}^{-1} \\ &- \frac{1}{2m_{\mathrm{P}}^2}\mathcal{G}^{mn}(\mathcal{D}_m\mathcal{D}_n\boldsymbol{P})\boldsymbol{P}^{-1} \\ &- \frac{1}{m_{\mathrm{P}}^2}\mathcal{G}^{mn}(\mathcal{D}_m\boldsymbol{P})\boldsymbol{P}^{-1}\,(\mathcal{D}_n\hat{\boldsymbol{U}}_0)\hat{\boldsymbol{U}}_0^{-1}.\end{aligned} \tag{5.191}$$

Here we have used the new notation

$$\mathcal{G}^{mn} = NG_{\perp}^{mn} \tag{5.192}$$

for another metric on the configuration space (cf. (5.168)), in which we have used the actual value of the lapse function corresponding to the classical extremal (5.185). This lapse function generally differs from unity. We have also decomposed the first-order corrections in the effective Hamiltonian into three terms, corresponding to the contribution of quantum matter (generated by $\hat{\boldsymbol{U}}_0$), a purely quantum-gravitational contribution generated by $\boldsymbol{P}$, and their cross-term.

Further evaluation of these terms demands a knowledge of derivatives acting on the configuration space argument q_+ of $\hat{\boldsymbol{U}}_0$ and $\boldsymbol{P}$. Obtaining these derivatives leads to the necessity of considering the special boundary value problem for classical equations of motion, the graviton propagator, and vertices—elements of the gravitational Feynman diagrammatic technique. The discussion is long and technical and can be found in Barvinsky and Kiefer (1998). Here we shall only quote the main steps and include a brief discussion of the results.

The following quantities play a role in the discussion. First, we introduce a collective notation for the full set of *Lagrangian* gravitational variables, which includes both the spatial metric and the lapse and shift functions,

$$g^a := (q^i(t), N^\mu(t)). \tag{5.193}$$

This comprises the space–time metric. (Recall that $q^i(t)$ stands for the three-metric $h_{ab}(\mathbf{x}, t)$.) Next, the second functional derivatives of the gravitational action with respect to the space–time metric are denoted by

$$S_{ab} := \frac{\delta^2 S[g]}{\delta g^a(t)\delta g^b(t')}. \tag{5.194}$$

Since the S_{ab} are not invertible, one must add gauge-fixing terms similar to (5.178). This leads to an operator F_{ab}. The 'graviton propagator' D^{bc} is then defined as its inverse via

$$F_{ab}D^{bc} = \delta_a^c. \tag{5.195}$$

We also need the components of the Wronskian operator obtained from the gravitational Lagrangian L^{g},

$$\overrightarrow{W}_{ib}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)\delta g^b(t) = -\delta\,\frac{\partial L^{\mathrm{g}}(q,\dot{q},N)}{\partial \dot{q}^i}. \tag{5.196}$$

With the help of the 'graviton propagator', one can define

$$\hat{t}^a(t) = -\frac{1}{m_{\mathrm{P}}^2}\int_{t_-}^{t_+}\mathrm{d}t'\, D^{ab}(t,t')\hat{T}_b(t') \equiv -\frac{1}{m_{\mathrm{P}}^2}\, D^{ab}\hat{T}_b, \tag{5.197}$$

where $\hat{T}_b$ is the condensed notation for the energy–momentum tensor of the matter field. The quantity $\hat{t}^a(t)$ obeys the linearized Einstein equations with source $\hat{T}_b$ and can thus be interpreted as the gravitational potential generated by the back reaction of quantum matter on the gravitational background.

The first correction term in (5.191)—the contribution of quantum matter—is found to read

$$\begin{aligned}
&-\frac{1}{2m_{\mathrm{P}}^2}\mathcal{G}^{mn}(\mathcal{D}_m\mathcal{D}_n\hat{\boldsymbol{U}}_0)\hat{\boldsymbol{U}}_0{}^{-1}\\
&= \frac{1}{2}m_{\mathrm{P}}^2\,\mathcal{G}^{mn}\,T\left(\overrightarrow{W}_{ma}\,\hat{t}^a\;\;\overrightarrow{W}_{nb}\,\hat{t}^b\right)\\
&\quad+\frac{\mathrm{i}}{2}D^{ab}w_{abc}(t_+)\,\hat{t}^c\\
&\quad-\frac{\mathrm{i}}{2}\mathcal{G}^{mn}\,(\overrightarrow{W}_{ma}D^{ac})\,(\overrightarrow{W}_{nb}D^{bd})\left(S_{cde}\,\hat{t}^e+\frac{1}{m_{\mathrm{P}}^2}\hat{S}^{\mathrm{mat}}_{cd}\right).
\end{aligned} \tag{5.198}$$

The resulting three terms can be given a Feynman diagrammatic representation, with different structures. Note that because of (5.197), all terms are of the same order m_{P}^{-2}, despite their appearance. The first term begins with the tree-level structure quadratic in the gravitational potential operators $\hat{t}^a$. Note that despite the fact that these operators are taken at one moment of time t_+, their chronological product is non-trivial because it should read as

$$T\left(\hat{t}^a\,\hat{t}^b\right) = \frac{1}{m_{\mathrm{P}}^4}\,D^{ac}D^{bd}\,T\left(\hat{T}_c\,\hat{T}_d\right) \tag{5.199}$$

and, thus, includes all higher-order chronological couplings between composite operators of matter stress tensors. The second and third terms on the right-hand side of (5.198) are essentially quantum, because their semiclassical expansions start with the one-loop diagrams consisting of one and two 'graviton propagators' D^{ab}, respectively. The quasi-local vertices of these diagrams are built from the Wronskian operators, gravitational three-vertices denoted by $w_{abc}(t_+)$ and S_{cde} (third functional derivatives of the gravitational action), and the second-order variation of the matter action with respect to the gravitational variables, $\hat{S}^{\mathrm{m}}_{cd}$. The corresponding diagrams are shown in Fig. 5.1.

The second and third correction terms in (5.191) can be written in a similar way. The second term—the purely quantum-gravitational contribution—contains instead of $\hat{t}^a$ a gravitational potential that is generated by the one-loop stress tensor of gravitons, which enters as a matter source in the linearized Einstein equations.

Depending on the physical situation, not all of the correction terms are of equal importance. It often happens that the effects of quantum matter dominate over the graviton effects and that, because of its contribution, only the first term on the right-hand side of (5.198) is significant. This means that one has

$$\hat{H}_1^{\mathrm{eff}} = \hat{H}^{\mathrm{m}} + \tfrac{1}{2}m_{\mathrm{P}}^2\,\mathcal{G}^{mn}\,\mathrm{T}\left(\overrightarrow{W}_{ma}\,\hat{t}^a\;\;\overrightarrow{W}_{nb}\,\hat{t}^b\right) + O(1/m_{\mathrm{P}}^2). \tag{5.200}$$

It turns out that this remaining correction term can be interpreted as the kinetic energy of the gravitational radiation produced by the back reaction of quantum matter sources. This term can be decomposed into a component along the semiclassical gravitational trajectory (the 'longitudinal part') and an orthogonal component; cf. Fig 4.3, in which the longitudinal part is called vertical part, and the orthogonal component is called horizontal part. The longitudinal part is ultralocal and basically given by the square of the matter Hamiltonian,

$$\propto \frac{1}{m_{\mathrm{P}}^2}\left(\hat{H}_{\perp}^{\mathrm{m}}\right)^2. \tag{5.201}$$

This is fully analogous to the relativistic correction term to the ordinary Schrödinger equation, as found from expanding the Klein–Gordon equation; see (5.149).

Can such quantum-gravitational correction terms be observed? If the matter Hamiltonian is dominated by the rest mass of a particle with mass m, the correction term (5.201) is of the order of

$$\propto \left(\frac{m}{m_{\mathrm{P}}}\right)^2. \tag{5.202}$$

For scalar fields causing an inflationary expansion of the early universe (inflatons), a typical value would be $m \approx 10^{-5}m_{\mathrm{P}}$. The corrections would thus be of the order 10^{-10}.

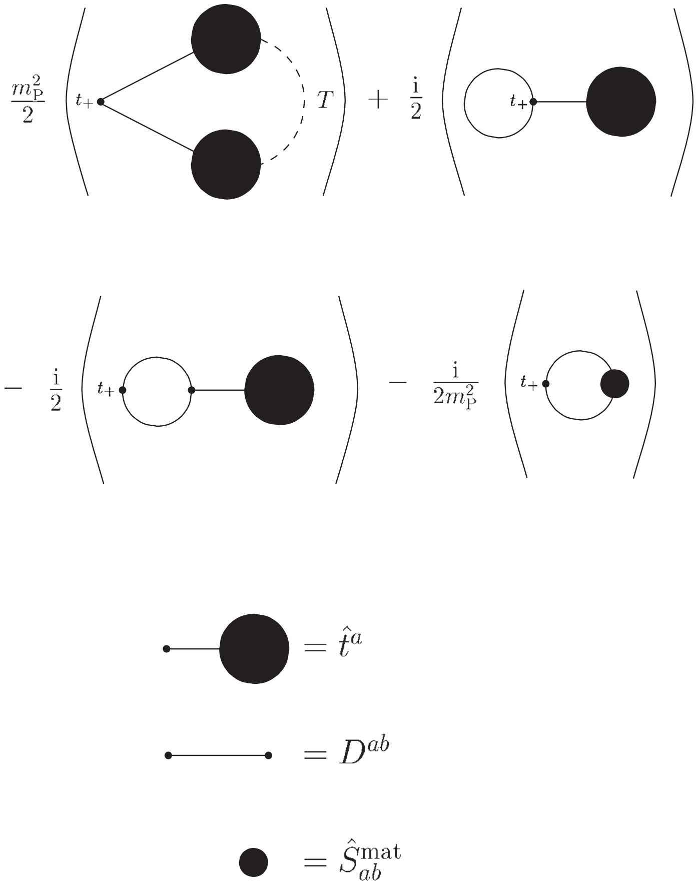

Fig. 5.1 Feynman diagrams illustrating the quantum-gravitational correction terms in (5.198). The time parameter t_+ labels the vertices at the space–time point at which $\hat{H}_1^{\mathrm{eff}}$ is evaluated. The dashed line labelled by T denotes the chronological ordering of matter stress tensors in the bilinear combinations of gravitational potentials $\hat{t}^a$. From Barvinsky and Kiefer 1998. © 1998 by Elsevier B.V.

More generally, one has instead of m the relevant energy scale of the non-gravitational degrees of freedom.

A simple example is the calculation of the quantum-gravitational correction to the trace anomaly in de Sitter space (Kiefer 1996). For a conformally coupled scalar field, the trace of the energy–momentum tensor, although it is zero classically, is non-vanishing in the quantum theory; this 'anomalous trace' is proportional to $\hbar$. It corresponds to the following expectation value ε of the Hamiltonian density,

$$\varepsilon = \frac{\hbar H_{\mathrm{dS}}^4}{1440\pi^2 c^3},$$

where H_{dS} denotes the constant Hubble parameter of de Sitter space. The first quantum-gravitational correction calculated from the Born–Oppenheimer expansion reads

$$\delta\varepsilon \approx -\frac{2G\hbar^2 H_{\mathrm{dS}}^6}{3(1440)^2\pi^3 c^8},$$

so that the ratio is given by

$$\frac{\delta\varepsilon}{\varepsilon} \approx -\frac{1}{2160\pi}\left(\frac{t_{\mathrm{P}}}{H_{\mathrm{dS}}^{-1}}\right)^2 . \tag{5.203}$$

One might perhaps have guessed from dimensional arguments that the ratio of the Planck time to the Hubble time would enter this expression, but the numerical prefactor can only be found from a concrete calculation. For realistic scenarios, the ratio (5.203) is, of course, small. Using values motivated by inflationary cosmology, one can assume that H_{dS} lies between 10^{13} and 10^{15} GeV, leading to values between roughly 10^{-16} and 10^{-22} for the ratio (5.203).

This raises the question of whether situations can be found that lead to observable effects. One possibility could be to look for quantum-gravitational contributions to the cosmic microwave background (CMB) anisotropies. The spectrum of these anisotropies has been measured with high precision by the WMAP satellite and other missions and is expected to be measured with even higher accuracy by the PLANCK mission. At large angular scales, the underlying primordial spectrum turns out to be approximately flat; that is, it is independent of the wave number k associated with a particular fluctuation. The contribution of the correction terms from (5.200) breaks this scale invariance (Kiefer and Krämer 2012). Denoting the undisturbed power spectrum of the fluctuations by $\Delta^2_{(0)}(k)$, one can write the corrected power spectrum in the form

$$\Delta^2_{(1)}(k) = \Delta^2_{(0)}(k)\, C_k^2 .$$

The correlation function C_k has been evaluated by Kiefer and Krämer (2012) for a particular model of chaotic inflation with a Hubble parameter H_{I} (cf. Linde 1990). The value $H_{\mathrm{I}} = 10^{14}$ GeV is a limiting value in the sense that existing observations constrain the Hubble parameter of inflation to be smaller that about 10^{14} GeV. For this limiting value, the corrections occur at such large scales (small k) that they are today beyond the cosmological horizon and thus unobservable. From the current non-observation one can, however, find an upper limit for H_{I} given by

$$H_{\mathrm{I}} \lesssim 4 \times 10^{17}\,\mathrm{GeV} . \tag{5.204}$$

Although this is a much weaker limit than the one already known, it is a definite prediction from a particular approach to quantum gravity. It is clear that the quantum-gravitational correction terms would become of order unity if H_{I} approached the Planck scale, but then the whole approximation scheme (and the very idea of inflation) would break down anyway.

The semiclassical approximation scheme has also been discussed for the constraint equations of supersymmetric canonical quantum gravity (Kiefer *et al.* 2005). It turns out that the formalism is only consistent if the states at each order depend on the gravitino field. That is, the Hamilton–Jacobi equation and therefore the background space–time must already involve the gravitino.

Further reading: DeWitt (1967), Isham (1993), Kuchař (1992), Moniz (2010).

6

Quantum gravity with connections and loops

6.1 Connection and loop variables

In Section 4.3, we encountered a Hamiltonian formulation of GR alternative to geometrodynamics, using the concepts of connections and loops. In the present chapter, we shall review approaches to formulating a consistent quantum theory with such variables, leading to 'quantum connection dynamics' or 'quantum loop dynamics'. More details and references can be found in Rovelli (2004), Thiemann (2007), Ashtekar and Lewandowski (2004), Gambini and Pullin (2011), and Nicolai *et al.* (2005). 'A new look at loop quantum gravity' was presented by Rovelli (2011). A short review is provided by Sahlmann (2011). The popular name 'loop quantum gravity' stems from the loop variables used in Section 6.1.2.

In the present section, we shall introduce the quantum versions of the connection and loop variables and discuss the Gauss constraints (4.124) and the diffeomorphism constraints (4.129). These constraints already give a picture of the way space might look like on the smallest scales. At least at a *kinematical* level (using the Gauss constraints only), a major result can be obtained: the spectrum of geometric operators representing area or volume in the classical limit turns out to be discrete (Section 6.2). This has a direct bearing on the interpretation of black-hole entropy (see Section 7.1). It is, of course, a long way from ensuring that this result holds also at the full dynamical level, where all constraints are solved.

The most complicated (and open) problem is the implementation of the Hamiltonian constraint; its status is discussed in Section 6.3. We treat here only pure gravity, but various results can be extended to the case of Standard Model matter (Thiemann 2007).

6.1.1 Connection representation

The connection representation is characterized classically by the Poisson bracket (4.122) between the densitized tetrad $E^b_j(\mathbf{y})$ and the SU(2) connection $A^i_a(\mathbf{x})$. These variables are then formally turned into operators obeying the commutation relation

$$\left[\hat{A}^i_a(\mathbf{x}), \hat{E}^b_j(\mathbf{y})\right] = 8\pi\beta\mathrm{i}\hbar\delta^i_j\delta^b_a\delta(\mathbf{x},\mathbf{y}). \tag{6.1}$$

In the functional Schrödinger representation, one can implement this relation formally through

$$\hat{A}_a^i(\mathbf{x})\Psi[A] = A_a^i(\mathbf{x})\Psi[A], \tag{6.2}$$

$$\hat{E}_j^b(\mathbf{y})\Psi[A] = 8\pi\beta\frac{\hbar}{\mathrm{i}}\frac{\delta}{\delta A_b^j(\mathbf{y})}\Psi[A], \tag{6.3}$$

where the A in the argument of the wave functional is a shorthand for $A_a^i(\mathbf{x})$. As in Chapter 5, the constraints are implemented as conditions on allowed wave functionals. The Gauss constraints (4.124) then become

$$\hat{\mathcal{G}}_i\Psi = 0 \longrightarrow \mathcal{D}_a\frac{\delta\Psi}{\delta A_a^i} = 0. \tag{6.4}$$

They express the invariance of the wave functional with respect to infinitesimal gauge transformations of the connection. The diffeomorphism constraints (4.129) become

$$\hat{\tilde{\mathcal{H}}}_a\Psi = 0 \longrightarrow F_{ab}^i\frac{\delta\Psi}{\delta A_b^i} = 0. \tag{6.5}$$

As in the classical case, they express the invariance of the wave functional under a combination of infinitesimal diffeomorphism and gauge transformations; cf. Section 4.3.2.

The Hamiltonian constraint (4.128) cannot be treated directly in this way, because the Γ_a^i terms contain the tetrad in a complicated non-linear fashion. This would lead to problems similar to those with the Wheeler–DeWitt equation discussed in Chapter 5, preventing one from finding any solutions. In Section 6.3, we shall see how a direct treatment of the quantum Hamiltonian constraint can at least be attempted. Here, we remark only that (4.128) is easy to handle only for the value $\beta = \mathrm{i}$ in the Lorentzian case and for the value $\beta = 1$ in the Euclidean case. In the first case, the problem arises that the resulting formalism uses complex variables and that one has to impose 'reality conditions' at an appropriate stage (which nobody has succeeded in doing). In the second case, the variables are real but one is dealing with the unphysical Euclidean case. Nevertheless, for these particular values of β, the second term in (4.128) vanishes, and the quantum Hamiltonian constraint for $\Lambda = 0$ simply reads

$$\epsilon^{ijk}F_{kab}\frac{\delta^2\Psi}{\delta A_a^i\delta A_b^j} = 0. \tag{6.6}$$

Note that in (6.4)–(6.6) a 'naive' factor ordering has been chosen: all derivatives are put to the right. Formal solutions to these equations have been found; see for example, Brügmann (1994). Some solutions have been expressed in terms of knot invariants.[1] Many of these solutions are annihilated by the operator corresponding to $\sqrt{h}$ and may therefore be devoid of physical meaning, since matter fields and the cosmological term couple to $\sqrt{h}$.

In the case of vacuum gravity with $\Lambda \neq 0$, an exact formal solution in the connection representation was found by Kodama (1990). Using a factor ordering different from (6.6), the Hamiltonian constraint reads for $\beta = 1$

[1] A knot invariant is a functional on the space of loops which assigns the same number to loops in the same knot class.

$$\epsilon^{ijk}\frac{\delta}{\delta A_a^i}\frac{\delta}{\delta A_b^j}\left(F_{kab}-\frac{\mathrm{i}\hbar\Lambda}{6}\epsilon_{abc}\frac{\delta}{\delta A_c^k}\right)\Psi[A]=0. \tag{6.7}$$

We note that the second term in parentheses comes from the determinant of the three-metric,

$$h=\det E_i^a=\frac{\mathrm{i}\hbar^3}{6}\epsilon^{ijk}\epsilon_{abc}\frac{\delta}{\delta A_a^i}\frac{\delta}{\delta A_b^j}\frac{\delta}{\delta A_c^k}.$$

The solution for Ψ is given by

$$\Psi_\Lambda=\exp\left(\mathrm{i}\frac{6}{G\Lambda\hbar}S_{\mathrm{CS}}[A]\right), \tag{6.8}$$

where $S_{\mathrm{CS}}[A]$ denotes the 'Chern–Simons action'

$$S_{\mathrm{CS}}[A]=\int_\Sigma \mathrm{d}^3x\ \epsilon^{abc}\,\mathrm{tr}\left(G^2A_a\partial_bA_c-\frac{2}{3}G^3A_aA_bA_c\right). \tag{6.9}$$

This follows after one notes that

$$\epsilon_{abc}\frac{\delta\Psi_\Lambda}{\delta A_c^k}=-\frac{6\mathrm{i}}{\Lambda\hbar}F_{kab}\ ;$$

that is, the term in parentheses in (6.7) by itself annihilates the state Ψ_Λ. Due to the topological nature of the Chern–Simons action, the state (6.8) is both gauge- and diffeomorphism-invariant. In contrast to the states mentioned above (for vanishing cosmological constant), (6.8) is not annihilated by the operator corresponding to $\sqrt{h}$ and may in this respect have physical content. The Chern–Simons action is also important for GR in 2+1 dimensions; cf. Section 8.1.3.

A state of the form (6.8) is also known from Yang–Mills theory. The state

$$\Psi_g=\exp\left(-\frac{1}{2\hbar g^2}S_{\mathrm{CS}}[A]\right)$$

is an eigenstate of the Yang–Mills Hamiltonian

$$H_{\mathrm{YM}}=\frac{1}{2}\int \mathrm{d}^3x\ \mathrm{tr}\left(-g^2\hbar^2\frac{\delta^2}{\delta A_a^2}+\frac{B_a^2}{g^2}\right),$$

where $B_a=(1/2)\epsilon_{abc}F^{bc}$, with eigenvalue zero (Loos 1969, Witten 2003). However, this state has unpleasant properties (e.g. it is not normalizable), which renders its physical significance dubious. It is definitely not the ground state of Yang–Mills theory. The same reservation may apply to (6.8).

Since the 'real' quantum Hamiltonian constraint is not given by (6.6) (see Section 6.3), we shall not discuss this type of solution further. What can generally be said about the connection representation? Since (6.4) guarantees that $\Psi[A]=\Psi[A^g]$, where $g\in\mathrm{SU}(2)$, the configuration space after the implementation of the Gauss constraints is actually given by $\mathcal{A}/\mathcal{G}$, where $\mathcal{A}$ denotes the space of connections and $\mathcal{G}$ the local

SU(2) gauge group. Because the remaining constraints have not been considered at this stage, this level corresponds to having, in Chapter 5, states $\Psi[h_{ab}]$ before imposing the constraints $\mathcal{H}_a\Psi = 0 = \mathcal{H}_\perp\Psi$. A candidate for an inner product on $\mathcal{A}/\mathcal{G}$ would be

$$\langle\Psi_1|\Psi_2\rangle = \int_{\mathcal{A}/\mathcal{G}} \mathcal{D}\mu[A]\ \Psi_1^*[A]\Psi_2[A]\ ; \tag{6.10}$$

cf. (5.20). The main problem is: can one construct a suitable measure $\mathcal{D}\mu[A]$ in a rigorous way? The obstacles are that the configuration space $\mathcal{A}/\mathcal{G}$ is both non-linear and infinite-dimensional. Such a measure has been constructed; see Ashtekar *et al.* (1994). In the construction process, it was necessary to extend the configuration space to its closure $\overline{\mathcal{A}/\mathcal{G}}$. This space is much bigger than the classical configuration space of smooth field configurations, since it contains distributional analogues of gauge-equivalent connections. The same should also hold for the wave functional $\Psi[h_{ab}]$ in quantum geometrodynamics (Isham 1976).

The occurrence of distributional configurations can also be understood from a path-integral point of view, where one sums over (mostly) non-differentiable configurations. In field theory, an imprint of this is left on the boundary configuration, which shows up as the argument of the wave functional. Although classical configurations form a set of measure zero in the space of all configurations, they nevertheless possess physical significance, since one can construct semiclassical states that are concentrated on them. Moreover, for the measurement of field variables, one would not expect much difference compared with the case of having smooth field configurations only, since only measurable functions count, and these are integrals of field configurations; see Bohr and Rosenfeld (1933). In a sense, the loop representation to be discussed in the following implements 'smeared versions' of the variables A and E.

6.1.2 Loop representation

Instead of considering wave functionals defined on the space of connections, $\Psi[A]$, one can use states defined on the space of *loops* $\alpha^a(s)$, $\Psi[\alpha]$; cf. Section 4.3.3. This is possible due to the availability of the measure on $\overline{\mathcal{A}/\mathcal{G}}$, and the states are obtained by the transformation ('loop transform')

$$\Psi[\alpha] = \int_{\overline{\mathcal{A}/\mathcal{G}}} \mathcal{D}\mu[A]\ \mathcal{T}[\alpha]\Psi[A], \tag{6.11}$$

where $\mathcal{T}[\alpha]$ was defined in (4.146). This corresponds to the usual Fourier transform in quantum mechanics,

$$\tilde{\psi}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{\mathbb{R}^3} \mathrm{d}^3x\ \mathrm{e}^{-\mathrm{i}\mathbf{px}/\hbar}\psi(\mathbf{x}). \tag{6.12}$$

A plane wave corresponds to $\mathcal{T}[\alpha] := \Psi_\alpha[A]$. We shall refer to the latter as 'loop states'. In the loop approach to quantum gravity, they can be taken to be the basis states (Rovelli and Smolin 1990). The prevalent opinion nowadays is that one should start directly with the loop variables, without reference to the connection representation and the loop transform.

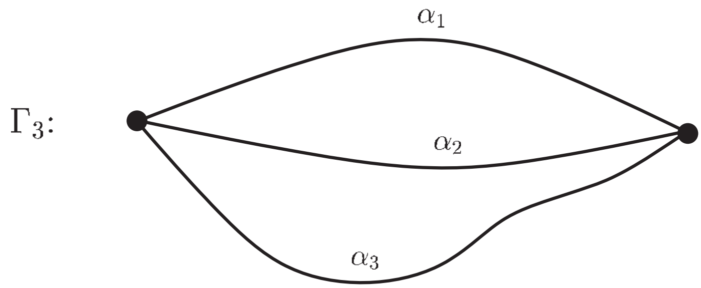

Fig. 6.1 An example of a graph with three curves.

The problem with the above loop states is that they form an overcomplete basis. A complete basis can be constructed by a linear combination of this basis. It is called the *spin-network basis* and goes back to Penrose (1971); see also, for example, Major (1999) for an introduction. To quote from Penrose (1971, p. 151):

My basic idea is to try and build up both space-time and quantum mechanics simultaneously—from *combinatorial* principles ... The idea here, then, is to start with the concept of angular momentum—where one has a *discrete* spectrum—and use the rules for combining angular momenta together and see if in some sense one can construct the concept of *space* from this.

How is the spin-network basis defined? Consider first a graph $\Gamma_n = \{\alpha_1, \ldots, \alpha_n\}$, where the α_i denote curves (also called 'edges' or 'links'), which are oriented and piecewise analytic. If they meet, they meet at their endpoints ('vertices' or 'nodes'). An example with three curves is depicted in Fig. 6.1. One then associates a holonomy $U[A, \alpha]$ (see (4.145)) to each link. This leads to the concept of 'cylindrical functions': considering a function

$$f_n : [\mathrm{SU}(2)]^n \longrightarrow \mathbb{C},$$

one can define the cylindrical function[2]

$$\Psi_{\Gamma_n, f_n}[A] = f_n(U_1, \ldots, U_n), \tag{6.13}$$

which depends on the connection through a finite but arbitrary number of holonomies. An SU(2) holonomy has thus been put on each of the n links of the graph.[3] Cylindrical functions are dense in the space of smooth functions on $\mathcal{A}$. One can define a scalar product between two cylindrical functions f and g that is invariant under gauge transformations and diffeomorphisms,

$$\langle \Psi_{\Gamma,f} | \Psi_{\Gamma,g} \rangle = \int_{[SU(2)]^n} \mathrm{d}U_1 \cdots \mathrm{d}U_n \; f^*(U_1, \ldots, U_n) g(U_1, \ldots, U_n), \tag{6.14}$$

where $\mathrm{d}U_1 \cdots \mathrm{d}U_n$ denotes the Haar measure. For different graphs, $\Gamma \neq \Gamma'$, one has $\langle \Psi_{\Gamma,f} | \Psi_{\Gamma',g} \rangle = 0$. With some assumptions, this scalar product is unique. It is basically

[2] The name 'cylindrical function' stems from the fact that these functions probe the connection A only along one-dimensional structures, that is, on a set of measure zero.

[3] It has also been suggested that one should use the group SO(3) instead of SU(2); cf. the discussion in Section 7.3 on black-hole entropy.

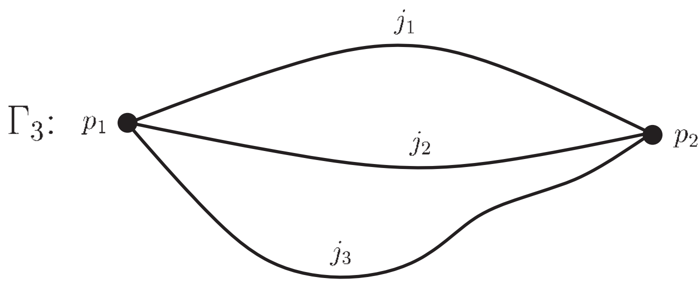

Fig. 6.2 The graph of Fig. 6.1 with spins attached to the links.

the choice of this scalar product that brings in the discrete structure of loop quantum gravity. Because the fundamental concepts here are graphs and spin networks instead of loops, the term 'quantum geometry' is sometimes used instead of loop quantum gravity; in order to avoid confusion with quantum geometrodynamics, however, we shall avoid this term.

At the beginning of Section 5.1, we distinguished between three spaces satisfying $\mathcal{F}_{\rm phys} \subset \mathcal{F}_0 \subset \mathcal{F}$, which do not necessarily have to be Hilbert spaces. Here, the intention is to use the Hilbert-space machinery of ordinary quantum theory as much as possible, and one would like to employ a chain of the form

$$\mathcal{H}_{\rm kin} \supset \mathcal{H}_{\rm g} \supset \mathcal{H}_{\rm diff} \supset \mathcal{H}_{\rm phys} \tag{6.15}$$

of Hilbert spaces in which the three sets of constraints (Gauss, diffeomorphism, and Hamiltonian constraints) are implemented consecutively. However, this would be possible only if the solutions to the constraints were normalizable. Since this is not the case, one has to employ the formalism of Gel'fand triples (rigged Hilbert spaces), which will not be elaborated on here; cf. Thiemann (2007).

The 'biggest' space, $\mathcal{H}_{\rm kin}$, is obtained by considering all linear combinations of cylindrical functions,

$$\Psi = \sum_{n=1}^{\infty} c_n \Psi_{\Gamma_n, f_n},$$

such that their norm is finite,

$$\parallel \Psi \parallel^2 = \sum_{n=1}^{\infty} |c_n|^2 \parallel \Psi_{\Gamma_n, f_n} \parallel^2 < \infty,$$

where

$$\parallel \Psi_{\Gamma,f} \parallel = \langle \Psi_{\Gamma,f} | \Psi_{\Gamma,f} \rangle^{1/2}.$$

The Hilbert space $\mathcal{H}_{\rm kin}$ itself is of course infinite-dimensional and carries unitary representations of local SU(2) transformations and diffeomorphisms. It is not separable; that is, it does not admit a countable basis.

With these preparations, a spin network is defined as follows. One associates with each link α_i a non-trivial irreducible representation of SU(2), that is, attaches a 'spin'

j_i to it ('colouring of the link'), where $j_i \in \{1/2, 1, 3/2, \ldots\}$. The representation acts on a Hilbert space $\mathcal{H}_{j_i}$. An example is shown in Fig. 6.2. Consider now a 'node' p where k links meet and associate to it the Hilbert space

$$\mathcal{H}_p = \mathcal{H}_{j_1} \otimes \ldots \otimes \mathcal{H}_{j_k}. \tag{6.16}$$

One fixes an orthonormal basis in $\mathcal{H}_p$ and calls an element of this basis a 'colouring' of the node p. A spin network is then a triple $S(\Gamma, \vec{j}, \vec{N})$, where $\vec{j}$ denotes the collection of spins and $\vec{N}$ the collection of basis elements at the nodes, that is, $\vec{N} = (N_{p_1}, N_{p_2}, \ldots)$, where N_{p_1} is a basis element at p_1, N_{p_2} a basis element at p_2, and so on. Note that S is not yet gauge-invariant. A *spin-network state* $\Psi_S[A]$ is then defined as a cylindrical function f_S associated with S. How is it constructed? One takes a holonomy at each link in the representation corresponding to j_i (described by 'matrices' $R^{j_i}(U_i)$) and contracts all these matrices with the chosen basis element $\in \mathcal{H}_p$ at each node where these links meet. This gives a complex number. Thus,

$$\Psi_S[A] = f_S(U_1, \ldots, U_n) = \prod_{\text{links } i} R^{j_i}(U_i) \otimes \prod_{\text{nodes } p} N_p, \tag{6.17}$$

where $\otimes$ refers here to the contraction of all indices: there is always one index of the matrix R (not indicated) that matches one index of a node basis element. One can prove that any two states Ψ_S are orthonormal:

$$\langle \Psi_S | \Psi_{S'} \rangle = \delta_{\Gamma\Gamma'} \delta_{\vec{j}\vec{j}'} \delta_{\vec{N},\vec{N}'} =: \delta_{SS'}. \tag{6.18}$$

The states Ψ_S form a complete basis in the unconstrained ('kinematical') Hilbert space $\mathcal{H}_{\text{kin}}$.

A gauge-invariant spin network can be constructed by imposing the Gauss constraints (6.4), leading to the Hilbert space $\mathcal{H}_{\text{g}}$. For this purpose, one first decomposes $\mathcal{H}_p$ into its irreducible parts ('Clebsch–Gordan decomposition'),

$$\mathcal{H}_p = \mathcal{H}_{j_1} \otimes \cdots \otimes \mathcal{H}_{j_k} = \bigoplus_J (\mathcal{H}_J)^{k_J}, \tag{6.19}$$

where k_J is the multiplicity of the spin-J irreducible representation. One then selects the *singlet* ($J = 0$) subspace, $(\mathcal{H}_0)^{k_0}$, which is gauge-invariant. One chooses at each node p an arbitrary basis and assigns one basis element to the node. The corresponding colouring $\vec{N}$ belongs to a gauge-invariant spin network.[4] At each node p, the spins of the meeting links have to obey the Clebsch–Gordan condition for any two pairs, for example $|j_1 - j_2| \leq j_3 \leq j_1 + j_2$, etc. for the situation of a 3-valent vertex shown in Fig. 6.3. In this example, there exists a unique intertwiner given by the Wigner $3j$ coefficient (there exists only one possible way to combine these three representations into a singlet). To give one example, considering the spin network of Fig. 6.2 with

[4]The colouring $\vec{N}$ is a collection of invariant tensors N_p, which are also called 'intertwiners' because they couple different representations of SU(2). They thus possess indices referring to different SU(2) representations.

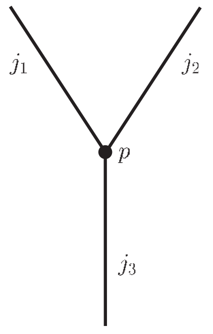

Fig. 6.3 Spins at meeting links have to obey Clebsch–Gordan conditions.

$j_1 = 1$, $j_2 = 1/2$, $j_3 = 1/2$, we find that the gauge-invariant spin-network state is given by the expression (Rovelli 2004, p. 236)

$$\Psi_S[A] = \frac{1}{3}\sigma_{i,AB} R^1(U[A,\alpha_1])^i{}_j U[A,\alpha_2]^A{}_C U[A,\alpha_3]^B{}_D \sigma^{j,CD},$$

where the σ_i are the Pauli matrices.

For nodes of valence four and higher, different choices are possible; cf. Nicolai *et al.* (2005). Several quantum states can thus be attributed to each spin network. We note that for a gauge-invariant spin network, angular momentum is conserved at each vertex. We also note that spin-network states can be decomposed into loop states; see Rovelli and Gaul (2000) for an illustrative example. In the literature, the kinematical Hilbert space is sometimes directly identified with $\mathcal{H}_{\mathrm{g}}$. For more details of the machinery of spin networks, we refer to Rovelli (2004).

The next step is the implementation of the diffeomorphism constraints (6.5). We shall denote the gauge-invariant spin-network states by

$$\Psi_S[A] := \langle A|S\rangle. \tag{6.20}$$

Diffeomorphisms move points on Σ around, so that the spin network will be 'smeared' over Σ. This leads to the concept of an 's-knot': two spin networks S and S' lie in the same s-knot if there exists a diffeomorphism $\phi \in \mathrm{Diff}\,\Sigma$ such that S' is the composition of ϕ with S, $S' = \phi \circ S$. One thus invokes a process of 'group averaging' instead of a direct application of the diffeomorphism constraint operator. This procedure throws the state out of the kinematical Hilbert space; therefore, the ideal relation $\mathcal{H}_{\mathrm{diff}} \subset \mathcal{H}_{\mathrm{kin}}$ does not hold and one must apply a more complicated construction involving rigged Hilbert spaces (Nicolai *et al.* 2005). In contrast to the situation for $\mathcal{H}_{\mathrm{kin}}$, it is expected that $\mathcal{H}_{\mathrm{diff}}$ will be separable ($\mathcal{H}_{\mathrm{diff}}$ is the Hilbert space for the averaged states). One defines

$$\langle s|S\rangle = \begin{cases} 0, \; S \notin s \;, \\ 1, \; S \in s \;, \end{cases} \tag{6.21}$$

and

$$\langle s|s'\rangle = \begin{cases} 0, & s \neq s' \;, \\ c(s), & s = s' \;, \end{cases} \tag{6.22}$$

where $c(s)$ denotes the number of discrete symmetries of the s-knots under diffeomorphisms (change of orientation and ordering). The diffeomorphism-invariant quantum states of the gravitational field are then denoted by $|s\rangle$. The important property is the non-local, 'smeared' character of these states, avoiding problems that such constructions would have, for example, in QCD. For details and references to the original literature, we refer to Thiemann (2007) and Rovelli (2004). The Hilbert space $\mathcal{H}_{\rm diff}$ is also used as the starting point for the action of the Hamiltonian constraint; see Section 6.3.

6.2 Quantization of area

Up to now, we have not considered operators acting on spin-network states. In the following, we shall construct one particular operator of central interest—the 'area operator'. It corresponds in the classical limit to the area of two-dimensional surfaces. Since the central role in the formalism is played by the algebra of angular momentum, this area operator will turn out to have a discrete spectrum in $\mathcal{H}_{\rm kin}$. The discussion can be given both within the connection representation and within the loop representation. We shall restrict ourselves to the latter.

Instead of (6.2), one can consider the operator corresponding to the holonomy $U[A,\alpha]$ (see Section 4.3.3) acting on spin-network states,

$$\hat{U}[A,\alpha]\Psi_S[A] = U[A,\alpha]\Psi_S[A]. \tag{6.23}$$

Instead of the operator acting in (6.3), which is an operator-valued distribution, it turns out to be more appropriate to consider a 'smeared' version in which (6.3) is integrated over a two-dimensional manifold $\mathcal{S}$ (not to be confused with the spin network S) embedded in Σ,

$$\hat{E}_i[\mathcal{S}] := -8\pi\beta\hbar{\rm i}\int_{\mathcal{S}} {\rm d}\sigma^1{\rm d}\sigma^2\; n_a(\vec{\sigma})\frac{\delta}{\delta A_a^i[\mathbf{x}(\vec{\sigma})]}, \tag{6.24}$$

where the embedding is given by $(\sigma^1,\sigma^2)\equiv\vec{\sigma}\mapsto x^a(\sigma^1,\sigma^2)$, and

$$n_a(\vec{\sigma}) = \epsilon_{abc}\frac{\partial x^b(\vec{\sigma})}{\partial\sigma^1}\frac{\partial x^c(\vec{\sigma})}{\partial\sigma^2} \tag{6.25}$$

is the usual vectorial hypersurface element. The operator defined in (6.24) corresponds to the flux of E^a_i through a two-dimensional surface. The canonical variables of loop quantum gravity are thus the holonomy and this flux. They obey the commutation relation

$$\left[\hat{U}[A,\alpha],\hat{E}_i[\mathcal{S}]\right] = {\rm i}l_{\rm P}^2\beta\iota(\alpha,\mathcal{S})U[\alpha_1,A]\tau_iU[\alpha_2,A],$$

where $\iota(\alpha,\mathcal{S}) = \pm1, 0$ is the 'intersection number', which depends on the orientation of α and $\mathcal{S}$. For these commutation relations, one can prove a theorem analogous to the Stone–von Neumann theorem in quantum mechanics: under some mild assumptions, this holonomy-flux representation is unique (Fleischhack 2009; Lewandowski *et al.* 2006). The kinematical structure in loop quantum gravity is thus essentially unique.

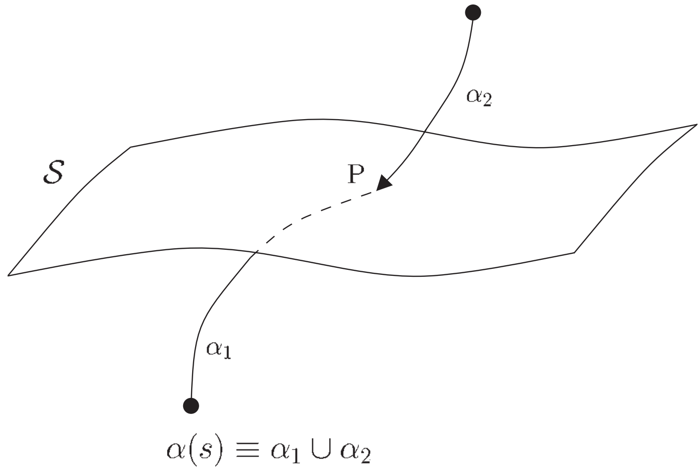

Fig. 6.4 Example of an intersection of a link α with a surface $\mathcal{S}$.

In this representation, there is no local operator that corresponds to the connection A_a^i itself.

We assume here the presence of only one intersection of α with $\mathcal{S}$; see Fig. 6.4, where α_1 refers to the part of α below $\mathcal{S}$ and α_2 to the part above $\mathcal{S}$. The intersection number vanishes if no intersection takes place. We want to emphasize that, in the following procedure, the diffeomorphism constraint has not yet been implemented.

We now want to calculate the action of $\hat{E}_i(\mathcal{S})$ on spin-network states $\Psi_S[A]$. For this, one needs its action on holonomies $U[A,\alpha]$. This was calculated in detail in Lewandowski *et al.* (1993) by using the differential equation (4.144) for the holonomy. In the simplest case of one intersection of α with $\mathcal{S}$ (as in Fig. 6.4), one obtains

$$\begin{aligned}\frac{\delta U[A,\alpha]}{\delta A_a^i[\mathbf{x}(\vec{\sigma})]} &= \frac{\delta}{\delta A_a^i[\mathbf{x}(\vec{\sigma})]}\left(\mathcal{P}\exp\left[G\int_\alpha \mathrm{d}s\ \dot{\alpha}^a A_a^i(\alpha(s))\tau_i\right]\right)\\ &= G\int_\alpha \mathrm{d}s\ \dot{\alpha}^a \delta^{(3)}(\mathbf{x}(\vec{\sigma}),\alpha(s))U[A,\alpha_1]\tau_i U[A,\alpha_2]. \end{aligned} \tag{6.26}$$

One can now act with the operator $\hat{E}_i[\mathcal{S}]$, (6.24), on $U[A,\alpha]$. This yields

$$\begin{aligned}\hat{E}_i[\mathcal{S}]U[A,\alpha] &= -8\pi\beta\hbar\mathrm{i}\int_{\mathcal{S}} \mathrm{d}\sigma^1\mathrm{d}\sigma^2\ \epsilon_{abc}\frac{\partial x^a(\vec{\sigma})}{\partial\sigma^1}\frac{\partial x^b(\vec{\sigma})}{\partial\sigma^2}\frac{\delta U[A,\alpha]}{\delta A_c^i[\mathbf{x}(\vec{\sigma})]}\\ &= -8\pi\beta\hbar G\mathrm{i}\int_{\mathcal{S}} \mathrm{d}\sigma^1\mathrm{d}\sigma^2 \int_\alpha \mathrm{d}s\ \epsilon_{abc}\frac{\partial x^a(\vec{\sigma})}{\partial\sigma^1}\frac{\partial x^b(\vec{\sigma})}{\partial\sigma^2}\frac{\partial\alpha^c}{\partial s}\\ &\quad\times\delta^{(3)}(\mathbf{x}(\vec{\sigma}),\alpha(s))U[A,\alpha_1]\tau_i U[A,\alpha_2].\end{aligned}$$

The coordinate transformation $(\sigma^1,\sigma^2,s)\mapsto(x^1,x^2,x^3)$ with $\alpha := (0,0,x^3)$ leads to the Jacobian

$$J := \frac{\partial(\sigma^1,\sigma^2,s)}{\partial(x^1,x^2,x^3)} = \epsilon_{abc}\frac{\partial x^a}{\partial\sigma^1}\frac{\partial x^b}{\partial\sigma^2}\frac{\partial\alpha^c}{\partial s}$$

(note that the right-hand side is zero for curves lying within the surface $\mathcal{S}$). Therefore,

$$\begin{aligned}\int_{\mathcal{S}}\int_{\alpha} \mathrm{d}\sigma^1 \mathrm{d}\sigma^2 \mathrm{d}s\ \epsilon_{abc} \frac{\partial x^a(\vec{\sigma})}{\partial \sigma^1} \frac{\partial x^b(\vec{\sigma})}{\partial \sigma^2} \frac{\partial x^c(s)}{\partial s} \delta^{(3)}(\mathbf{x}(\vec{\sigma}), \mathbf{x}(s)) \\ = \int \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3\ \delta^{(3)}(\mathbf{x}(\vec{\sigma}), \mathbf{x}(s)) = \pm 1,\end{aligned} \tag{6.27}$$

where the sign depends on the relative orientation of the curve and the surface. The holonomy (defined on a one-dimensional edge) and the flux (defined on a two-dimensional surface) together yield the prerequisites for an integral over the three-dimensional delta function. One thus gets

$$\hat{E}_i[\mathcal{S}]U[A, \alpha] = \pm 8\pi\beta\hbar G \mathrm{i} U[A, \alpha_1]\tau_i U[A, \alpha_2]. \tag{6.28}$$

If there is no intersection, the action of this operator is zero. For more than one point of intersection, one has to sum over all of them.

What, then, is the action of $\hat{E}_i[\mathcal{S}]$ on a spin network? Consider a gauge-invariant spin network S intersecting $\mathcal{S}$ at a single point P. Then, decompose $\Psi_S[A]$ (cf. (6.17)) as

$$\Psi_S[A] = \Psi^{mn}_{S-\alpha}[A] R^j_{mn}(U[A, \alpha]) \,, \tag{6.29}$$

where $R^j_{mn}(U[A, \alpha])$ is the holonomy along α in the irreducible representation corresponding to spin j, and $\Psi^{mn}_{S-\alpha}[A]$ is the remaining part of (6.17). The action of $\hat{E}_i[\mathcal{S}]$ on R^j is similar to (6.28), with $\tau_i \to \tau_i^{(j)}$ according to the representation associated with j. Then,

$$\begin{aligned}\hat{E}_i[\mathcal{S}]\Psi_S[A] = \pm 8\pi\beta l_{\mathrm{P}}^2 \mathrm{i} \left[R^j(U[A, \alpha_1])\, \tau_i^{(j)} R^j(U[A, \alpha_2])\right]_{mn} \\ \times \Psi^{mn}_{S-\alpha}[A].\end{aligned} \tag{6.30}$$

This action is not yet gauge-invariant. One can obtain a gauge-invariant operator by 'squaring', that is, by considering

$$\hat{E}^2[\mathcal{S}] := \hat{E}_i[\mathcal{S}]\hat{E}_i[\mathcal{S}]. \tag{6.31}$$

In order to calculate the action of this operator, consider again a spin network with a single point of intersection P. We assume that P belongs to the α_1 part of the curve. Therefore, in the action

$$\begin{aligned}\hat{E}^2[\mathcal{S}]\Psi_S[A] = \pm 8\pi\beta l_{\mathrm{P}}^2 \mathrm{i} \hat{E}_i[\mathcal{S}] \left[R^j(U[A, \alpha_1])\, \tau_i^{(j)} R^j(U[A, \alpha_2])\right]_{mn} \\ \times \Psi^{mn}_{S-\alpha}[A],\end{aligned}$$

the operator $\hat{E}_i[\mathcal{S}]$ on the right-hand side acts only on $R^j(U[A, \alpha_1])$ to give $R^j(U[A, \alpha_1])\, \tau_i^{(j)} \mathbb{I}$, where $\mathbb{I}$ is the unit operator. Since one has for the 'Casimir operator'

$$\tau_i^{(j)} \tau_i^{(j)} = -j(j+1)\mathbb{I}$$

(recall that we have defined $\tau_i = (\mathrm{i}/2)\sigma_i$ for $j = 1/2$ and similarly for higher j), we get

$$\hat{E}^2[\mathcal{S}]\Psi_S[A] = (8\pi\beta l_{\mathrm{P}}^2)^2 j(j+1)\Psi_S[A], \tag{6.32}$$

where (6.17) has been used. If there is more than one intersection of S with $\mathcal{S}$, one has to consider a partition ρ of $\mathcal{S}$ into $n(\rho)$ smaller surfaces $\mathcal{S}_n$ such that the points of intersection lie in different $\mathcal{S}_n$ (for a given S). Otherwise, the action of $\hat{E}^2[\mathcal{S}]$ will not be gauge-invariant (due to 'cross-terms' in $\hat{E}^2[\mathcal{S}]\Psi_S[A]$).

We now define the 'area operator' (see below for its interpretation)

$$\hat{\mathcal{A}}[\mathcal{S}] := \lim_{\rho\to\infty} \sum_{n(\rho)} \sqrt{\hat{E}_i[\mathcal{S}_n]\hat{E}_i[\mathcal{S}_n]}, \tag{6.33}$$

which is independent of ρ. If there are no nodes on $\mathcal{S}$ and only a finite number of intersections ('punctures' P), one obtains from (6.32) (Rovelli and Smolin 1995; Ashtekar and Lewandowski 1997)

$$\hat{\mathcal{A}}[\mathcal{S}]\Psi_S[A] = 8\pi\beta l_{\mathrm{P}}^2 \sum_{\mathrm{P}\in S\cap\mathcal{S}} \sqrt{j_{\mathrm{P}}(j_{\mathrm{P}}+1)}\Psi_S[A] =: A[\mathcal{S}]\Psi_S[A]. \tag{6.34}$$

The operator $\hat{A}(\mathcal{S})$ is self-adjoint in $\mathcal{H}_{\mathrm{kin}}$; that is, it is diagonal on spin-network states and is real on them. Spin-network states are thus eigenstates of the area operator. Its spectrum is discrete because the spin network has a discrete structure. This feature can be traced back to the compactness of the group SU(2) used in the formalism. The three-manifold Σ is, however, still present.

In Chapter 2, we have seen that various approaches (asymptotic safety and dynamical triangulation) predict that the effective number of space–time dimensions approaches two at small scales. This can also be inferred from the area spectrum (6.34) (Modesto 2009). Defining $\ell_j := \sqrt{j_{\mathrm{P}}}l_{\mathrm{P}}$, one gets from (6.34)

$$A[\mathcal{S}] = 8\pi\beta \sum_{\mathrm{P}\in S\cap\mathcal{S}} \sqrt{\ell_j^2(\ell_j^2 + l_{\mathrm{P}}^2)},$$

which approaches ℓ_j^2 for large areas, but $\ell_j l_{\mathrm{P}}$ for small areas. One thus recognizes a different scaling with the length ℓ_j at small and large distances. Discussing the spectral dimension, Modesto (2009) finds that there is a decrease in the number of space dimensions from three at large scales to one at small scales.

If a node lies on $\mathcal{S}$, a more complicated expression is obtained (Frittelli *et al.* 1996, Ashtekar and Lewandowski 1997): denoting the nodes by $\vec{j}_i = (j_i^u, j_i^d, j_i^t)$, $i = 1, \ldots, n$, where j_i^u denotes the colouring of the upper link, j_i^d the colouring of the lower link, and j_i^t the colouring of a link tangential to $\mathcal{S}$, one obtains

$$\hat{\mathcal{A}}[\mathcal{S}]\Psi_S[A] = 4\pi\beta l_{\mathrm{P}}^2 \sum_{i=1}^{n} \sqrt{2j_i^u(j_i^u+1) + 2j_i^d(j_i^d+1) - j_i^t(j_i^t+1)}\Psi_S[A]. \tag{6.35}$$

For the special case $j_i^t = 0$ and $j_i^u = j_i^d$, we obtain again the earlier result (6.34). It must, however, be emphasized that spin-network states for 4-valent (and higher) vertices are not always eigenstates of $\hat{\mathcal{A}}$ if the node is on the surface, cf. Nicolai *et al.* (2005).

We now show that $\hat{\mathcal{A}}[\mathcal{S}]$ is indeed an 'area operator'; that is, the classical version of (6.33) is just the classical area of $\mathcal{S}$. This classical version is

$$E_i[\mathcal{S}_n] = \int_{\mathcal{S}_n} \mathrm{d}\sigma^1 \mathrm{d}\sigma^2 \; n_a(\vec{\sigma}) E_i^a(\mathbf{x}(\vec{\sigma})) \approx \Delta\sigma^1 \Delta\sigma^2 n_a(\vec{\sigma}) E_i^a(\mathbf{x}_n(\vec{\sigma})),$$

where $\mathcal{S}_n$ refers to a partition of $\mathcal{S}$, and $\mathbf{x}_n(\vec{\sigma})$ is an arbitrary point in $\mathcal{S}_n$ (it is assumed that the partition is sufficiently fine-grained). For the area, that is, the classical version of (6.33), one then obtains

$$\begin{aligned} A[\mathcal{S}] &= \lim_{\rho\to\infty} \sum_{n(\rho)} \Delta\sigma^1 \Delta\sigma^2 \sqrt{n_a(\vec{\sigma}) E_i^a(\mathbf{x}_n(\vec{\sigma})) n_b(\vec{\sigma}) E_i^b(\mathbf{x}_n(\vec{\sigma}))} \\ &= \int_{\mathcal{S}} \mathrm{d}^2\sigma \; \sqrt{n_a(\vec{\sigma}) E_i^a(\mathbf{x}_n(\vec{\sigma})) n_b(\vec{\sigma}) E_i^b(\mathbf{x}_n(\vec{\sigma}))}. \end{aligned}$$

Adopting coordinates on $\mathcal{S}$ as $x^3(\vec{\sigma}) = 0$, $x^1(\vec{\sigma}) = \sigma^1$, $x^2(\vec{\sigma}) = \sigma^2$, one gets from (6.25) that $n_1 = n_2 = 0$ and $n_3 = 1$. Using in addition (4.104) and (4.106), one obtains for the area

$$\begin{aligned} A[\mathcal{S}] &= \int_{\mathcal{S}} \mathrm{d}^2\sigma \; \sqrt{h(\mathbf{x}) h^{33}(\mathbf{x})} = \int_{\mathcal{S}} \mathrm{d}^2\sigma \; \sqrt{h_{11}h_{22} - h_{12}h_{21}} \\ &= \int_{\mathcal{S}} \mathrm{d}^2\sigma \; \sqrt{{}^{(2)}h}, \end{aligned} \tag{6.36}$$

where ${}^{(2)}h$ denotes the determinant of the two-dimensional metric on $\mathcal{S}$. Therefore, the results (6.34) and (6.35) demonstrate that *area is quantized* in units proportional to the Planck area l_{P}^2. A similar result holds for volume and length, although the discussion in this case is much more involved (cf. Thiemann 2007). It seems that the area operator is somehow distinguished; this could point to a formulation in terms of an 'area metric' that is usually discussed in another context; cf. Schuller and Wohlfarth (2006). We should also mention that, even at the kinematical level, a discrete spectrum does not necessarily follow in all dimensions: for example, Freidel *et al.* (2003) find that a space-like length operator in $2 + 1$ dimensions has a continuous spectrum.

The occurrence of discrete spectra for geometric operators might be an indication of the discreteness of space at the Planck scale, already mentioned in Chapter 1. Since these geometric quantities refer to three-dimensional space, they indicate only a discrete nature of space, not space–time. In fact, as we have seen in Section 5.4, space–time as a whole emerges only in a semiclassical limit. In spite of these results, the three-dimensional manifold Σ remains in the formalism of loop quantum gravity and thus still plays the role of an 'absolute' structure; cf. Section 1.3.

The discrete spectrum (6.34) and (6.35) is considered as one of the central results of quantum loop (or quantum connection) kinematics. Whether all eigenvalues of (6.35) are indeed realized depends on the topology of $\mathcal{S}$ (Ashtekar and Lewandowski 1997). In the case of an open $\mathcal{S}$ whose closure is contained in Σ, they are all realized. This is not the case for a closed surface.

The smallest eigenvalue of the area operator is zero. Its smallest non-zero eigenvalue is (in the case of open $\mathcal{S}$)[5]

$$A_0 = 2\pi\sqrt{3}\beta l_{\mathrm{P}}^2 \approx \beta \cdot 2.86 \times 10^{-65}\,\mathrm{cm}^2, \tag{6.37}$$

which is obtained from (6.35) for $j^d = 0$ and $j^u = j^t = j = 1/2$. The area gap (6.37) is referred to as one 'quantum of area'. Ashtekar and Lewandowski (1997) also showed that for $A[\mathcal{S}] \to \infty$, the difference ΔA between an eigenvalue A and its closest eigenvalue obeys

$$\Delta A \leq 4\pi\beta l_{\mathrm{P}}^2 \frac{\sqrt{8\pi\beta}}{\sqrt{A}} + \mathcal{O}\left(\frac{l_{\mathrm{P}}^2}{A}\right) l_{\mathrm{P}}^2. \tag{6.38}$$

Therefore, $\Delta A \to 0$ for large A.

Although the area operator $\hat{\mathcal{A}}(\mathcal{S})$ is gauge-invariant (invariant under SU(2) or SO(3) transformations), it is *not* invariant under three-dimensional diffeomorphisms. The reason is that it is defined for an abstract surface in terms of coordinates. It is thus not defined on $\mathcal{H}_{\mathrm{diff}}$. It also does not commute with the quantum Hamiltonian constraint. For the same reason, it is *not* an observable in the sense of Section 3.5. It is sometimes believed that it will become diffeomorphism-invariant (and an observable) if the surface is defined *intrinsically* through curvature invariants of the gravitational field or concrete matter fields; see, for example, Rovelli (2004) and the remarks in Section 4.2.3. But an explicit and complete construction in this spirit is elusive.

It is far from clear whether the discreteness of geometrical operators continues to hold if these operators are constructed in such a way that they commute with all constraints. Simple examples indicate that this will not be the case (Dittrich and Thiemann 2009).[6] It is also far from clear whether the area spectrum is of any operational significance in the sense of a measurement analysis with rods and clocks. This can only be decided after an analysis similar to that presented in Section 1.2 has been applied to the full theory.

It is somewhat surprising that an important issue such as the fundamental discreteness of space emerges even at the kinematical level. One would have instead expected that it would be a result that emerges from the treatment of the Hamiltonian constraint (Section 6.3), which encodes the 'dynamical' features of Einstein's theory. The discreteness thus seems to hold for more general theories than quantum general relativity.

The discrete spectrum of the area operator is also at the heart of the statistical foundation of black-hole entropy, that is, the recovery of the black-hole entropy through a quantitative counting of microscopic states for the gravitational field. This will be discussed in Section 7.3.

6.3 Quantum Hamiltonian constraint

The next task in the quantization process is the treatment of the Hamiltonian constraint. We recall that the Hilbert space $\mathcal{H}_{\mathrm{diff}}$ has not been obtained by the action of

[5]Such a gap between zero and the first non-vanishing eigenvalue does not seem to exist for the volume operator.

[6]But see the different opinion expressed in Rovelli (2007).

an operator, but instead with the help of a group averaging procedure. On the other hand, the Hamiltonian constraint is still being imposed as an operator constraint. The exact treatment of this constraint is the central (as yet open) problem in loop quantum gravity.

One of the main problems is that the Hamiltonian constraint operator does not preserve the Hilbert space $\mathcal{H}_{\rm diff}$. In order to guarantee diffeomorphism invariance, a more involved procedure must be invoked; cf. Nicolai *et al.* (2005). One can address the action of the constraint operator $\hat{\mathcal{H}}_\perp$ on a space $\Upsilon^* \subset S^*$, where S^* is the space dual to the spin-network space S. One can choose $\Upsilon^* = \mathcal{H}_{\rm diff}$, which, to emphasize the point again, is *not* a subset of $\mathcal{H}_{\rm kin}$. The central object is then a 'dual operator' $\hat{\mathcal{H}}^*_\perp$ that acts on Υ^*. This operator must first be regularized; see below. The limit of the corresponding regularization parameter going to zero can then only be implemented as a weak limit in the sense of matrix elements.

How is the Hamiltonian constraint dealt with concretely? In Section 4.3, we considered the rescaled constraint

$$\tilde{\mathcal{H}}_\perp = -8\pi G\beta^2 \mathcal{H}^{\rm g}_\perp \ ;$$

see (4.128) and (4.131). It can be written as

$$\begin{aligned}\tilde{\mathcal{H}}_\perp &= \frac{1}{\sqrt{h}}\mathrm{tr}\left(\left(F_{ab} + \frac{\beta^2\sigma - 1}{\beta^2}R_{ab}\right)[E^a, E^b]\right)\\ &=: \mathcal{H}_{\rm E} + \frac{\beta^2\sigma - 1}{\beta^2\sqrt{h}}\mathrm{tr}\left(R_{ab}[E^a, E^b]\right).\end{aligned} \tag{6.39}$$

The central idea in the quantization of this constraint is to use the fact that area and volume operators can be rigorously defined. (For the area operator, this was discussed in the last section. The volume operator can be treated analogously; see Ashtekar and Lewandowski 1998) A direct replacement of E^a by a derivative operator as in (6.3) would not lead very far, since R_{ab} depends on E^a in a complicated way. In this respect, the situation has not improved compared with the geometrodynamic approach discussed in Chapter 5.

In the first step, one addresses the 'Euclidean' part $\mathcal{H}_{\rm E}$ of (6.39).[7] One recognizes from (4.138) that this part depends solely on the volume, V, and on F_{ab} and A_c. The volume operator—the quantum equivalent to (4.136)—is well defined and yields a self-adjoint operator on the Hilbert space $\mathcal{H}_{\rm kin}$ with a finite action on cylindrical functions. There are therefore no problems with factor ordering at this stage. The operators corresponding to F_{ab} and A_c can be treated by using holonomies. In the second step, one makes use of (4.142), which in the Lorentzian case $\sigma = -1$ reads

$$\tilde{\mathcal{H}}_\perp + \frac{\mathcal{H}_E}{\beta^2} = -\frac{\beta^2+1}{2(4\pi\beta)^3}\epsilon^{abc}\,\mathrm{tr}\left(\{A_a, T\}\{A_b, T\}\{A_c, V\}\right). \tag{6.40}$$

Using (4.140), one can quantize T and then obtain a quantization of the full operator $\tilde{\mathcal{H}}_\perp$.

[7]In fact, the remaining part is rarely discussed, because its discussion is much more involved.

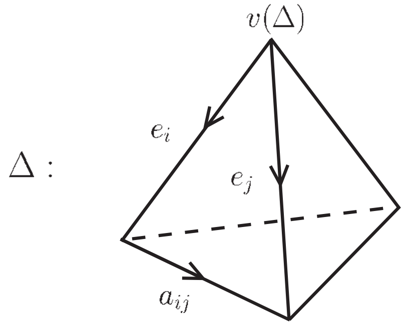

Fig. 6.5 An elementary tetrahedron Δ.

In order to get a well-defined operator $\mathcal{H}_{\mathrm{E}}$, Thiemann (1996) considered the integral of (4.138) with respect to a lapse function $N(\mathbf{x})$,

$$H_{\mathrm{E}}[N] = -\frac{1}{4\pi\beta} \int_\Sigma \mathrm{d}^3x \; N(\mathbf{x})\epsilon^{abc} \,\mathrm{tr}\,(F_{ab}\{A_c, V\})\,. \tag{6.41}$$

One can perform a triangulation (denoted by 'Tri') of Σ into elementary tetrahedra Δ and choose one vertex, $v(\Delta)$, for each Δ; see Fig. 6.5. If the three edges of Δ meeting at $v(\Delta)$ are denoted by $e_i(\Delta)$, $i = 1, 2, 3$, one can consider the loop

$$\alpha_{ij}(\Delta) := e_j^{-1}(\Delta) \circ a_{ij}(\Delta) \circ e_i(\Delta),$$

where $a_{ij}(\Delta)$ connects the vertices other than $v(\Delta)$. Thiemann (1996) could then show that one can get the correct Euclidean Hamiltonian (6.41) in the limit where all tetrahedra Δ shrink to their base points $v(\Delta)$. Consider

$$H_{\mathrm{E}}^{\mathrm{Tri}}[N] = \sum_{\Delta\in\mathrm{Tri}} H_{\mathrm{E}}^{\Delta}[N], \tag{6.42}$$

with

$$H_{\mathrm{E}}^{\Delta}[N] := \frac{1}{12\pi\beta} N\left(v(\Delta)\right) \epsilon^{ijk} \,\mathrm{tr}\left(U_{\alpha_{ij}(\Delta)} U_{e_k(\Delta)} \{U^{-1}_{e_k(\Delta)}, V\}\right), \tag{6.43}$$

where $U_{\dots}$ denotes the holonomies along the corresponding loops and edges. Using

$$\lim_{\Delta\to v(\Delta)} U_{\alpha_{ij}(\Delta)} = 1 + \tfrac{1}{2}\, F_{ab} e_i^a(\Delta) e_j^b(\Delta),$$
$$\lim_{\Delta\to v(\Delta)} U_{e_k(\Delta)} = 1 + A_a e_k^a(\Delta),$$

the expression (6.42) tends to (6.41) in this limit.

The quantum operator corresponding to (6.42) is then *defined* by replacing V by $\hat{V}$ and by replacing Poisson brackets with commutators,

$$\hat{H}_{\mathrm{E}}^{\mathrm{Tri}}[N] := \sum_{\Delta\in\mathrm{Tri}} \hat{H}_{\mathrm{E}}^{\Delta}[N], \tag{6.44}$$

where

$$\hat{H}_{\mathrm{E}}^{\Delta}[N] := -\frac{\mathrm{i}}{12\pi\beta\hbar} N\left(v(\Delta)\right) \epsilon^{ijk} \operatorname{tr}\left(\hat{U}_{\alpha_{ij}(\Delta)} \hat{U}_{e_k(\Delta)} \{\hat{U}^{-1}_{e_k(\Delta)}, V\}\right). \tag{6.45}$$

Furthermore, one can show that

$$\hat{H}_{\mathrm{E}}^{\mathrm{Tri}}[N] f_\alpha = \sum_{\Delta \in \mathrm{Tri}; \Delta \cap \alpha \neq \emptyset} \hat{H}_{\mathrm{E}}^{\Delta}[N] f_\alpha, \tag{6.46}$$

where f_α denotes a cylindrical function associated with a graph α. From inspection of the right-hand side, one recognizes that there is a contribution only if Δ intersects α (in fact, as one can show, only if it intersects it at a vertex). This gives only a finite number of such terms, yielding an 'automatic regularization' that survives in the 'continuum limit' $\Delta \to v(\Delta)$. Thiemann (1996) also showed that no anomalies arise in the constraint algebra 'on-shell'. Since the Hamiltonian constraint operator acts only on diffeomorphism-invariant states, the right-hand side of the quantum analogue to (3.84) vanishes, so nothing can be said about the full 'off-shell' algebra. He also showed that a Hermitian factor ordering can be chosen. When all this has been done, it is expected from (6.40) that a regularization can also be achieved for the full Hamiltonian constraint operator.

We also want to remark that it is unclear whether the operator $\hat{H}_{\mathrm{E}}$ has anything to do with the left-hand side of (6.6); the solutions mentioned there will most likely not be annihilated by $\hat{H}_{\mathrm{E}}$.

Various interesting open questions remain, which are the focus of current research. Among these are:

1. The definition of the Hamiltonian constraint is plagued with various ambiguities (operator ordering, choice of representation for the evaluation of the traces, ...); see Nicolai *et al.* (2005). There thus exists a whole class of different quantizations. Are there physical criteria which single out a unique operator?
2. Can physically interesting solutions to all constraints be obtained? This has at least been achieved for (2+1)-dimensional gravity and for cosmological models (see Thiemann 2007).
3. The action of the Hamiltonian constraint operator is different from the action of the Hamiltonian in lattice gauge theories. Analogies from there could therefore be misleading; see Nicolai *et al.* (2005).
4. Is the quantum constraint algebra anomaly-free 'off-shell', that is, before the constraints are implemented? As a comparison with the situation for the bosonic string (Section 3.2) demonstrates, insisting only on on-shell closure spoils the standard results of string quantization; see Nicolai *et al.* (2005).
5. Does one obtain the correct classical limit of the constraint algebra?
6. How does the semiclassical approximation scheme work?

The last question has not yet been addressed in the same way as was done with the geometrodynamical constraints in Section 5.4, since the constraints do not have a simple form in which a 'Born–Oppenheimer-type' method can be straightforwardly employed. Instead, researchers have tried to use the methods of coherent states; cf. Thiemann (2007). In order to avoid some of these problems, a 'master constraint programme' has been suggested (Thiemann 2006). There, one combines the smeared Hamiltonian constraints for all smearing functions into one single constraint. It is, however, premature to make any judgement about the success of this programme.

To summarize, the main results in the loop or connection representation have been found at the kinematical level. This holds in particular for the discrete spectra of geometric operators. The main open problem is the correct implementation (and solution) of the Hamiltonian constraint and the recovery of the classical limit.

Alternative approaches to the quantum Hamiltonian constraint are provided by 'spin-foam models', which use a path-integral type of approach employing the evolution of spin networks in 'time'. We shall not discuss this here, but refer the reader to the literature; see Rovelli (2004), Perez (2006), and Nicolai *et al.* (2005). Spin-foam models belong to a class of theories that approach the problem of quantum gravity from a fundamentally *discrete* perspective. These theories either have grown out of other approaches (such as loop quantum gravity or string theory) or attempt to tackle the problem using a new ansatz, as a 'secondary theory' in the sense of Isham (Section 1.3). A detailed overview of such theories can be found in the book edited by Oriti (2009). Examples are the theory of causal sets (or causet theory), topos theory, and geometrogenesis. All these approaches emphasize their background-independent nature. Group-field theory is a general framework that encompasses some of these discrete theories. It can be understood as a non-local quantum field theory for a field on a group manifold and corresponds conceptually to a third-quantized version of gravity. Some of these approaches exist so far only as classical discrete theories.

A major problem of many discrete approaches is the recovery of appropriate limits: this can be just the continuum limit, the limit of GR at the classical level, or—if the starting point is already a quantum theory—GR as the correct classical limit. There are, so far, few relations to phenomenology. One exception is the prediction of a small positive cosmological constant from causet theory (Ahmed *et al.* 2004). In fact, Λ is dynamical there and fluctuates around zero, but has a mean square deviation of the order of the currently observed dark energy.

Further reading: Ashtekar and Lewandowski (2004), Nicolai *et al.* (2005), Oriti (2009), Rovelli (2004), Thiemann (2007).

7
Quantization of black holes

7.1 Black-hole thermodynamics and Hawking radiation

In this section, we shall briefly review the thermodynamical behaviour of black holes and the Hawking effect. These issues arise at a semiclassical level—the gravitational field is treated as an external classical background—but they are expected to play a key role in the search for quantum gravity. More details can be found, for example, in Jacobson (2003), Kiefer (1998, 1999), and the references therein; see also the brief review by Kiefer (2003*a*), which we shall partially follow.

7.1.1 The laws of black-hole mechanics

It is a most amazing fact that black holes obey *uniqueness theorems* (see Heusler 1996 for a detailed exposition). If an object collapses to form a black hole, a stationary state will be reached asymptotically. One can prove within the Einstein–Maxwell theory that stationary black holes are uniquely characterized by only three parameters: their mass M, angular momentum $J \equiv Ma$, and electric charge q.[1] In this sense, black holes are much simpler objects than ordinary stars—given these parameters, they all look the same. All other degrees of freedom that might have initially been present have thus been radiated away, for example in the form of electromagnetic or gravitational radiation during the collapse, or have just disappeared (such as baryon number). Since these other degrees of freedom constitute some form of 'hair' (structure), this theorem is called the *no-hair theorem*. The three parameters are associated with conservation laws at spatial infinity. In principle, one can thus decide about the nature of a black hole far away from the hole itself, without having to approach it. In astrophysical situations, electrically charged black holes do not play an important role, so the two parameters M and J suffice. The corresponding solution of Einstein's equations is called the Kerr solution (one has the Kerr–Newman solution in the presence of charge). Stationary black holes are axially symmetric, with spherical symmetry being obtained as a special case for $J = 0$. Charged black holes are of interest for theoretical reasons. It should be emphasized that the uniqueness theorems do not generalize to higher dimensions in a straightforward way.

In the presence of other fields, the uniqueness theorems do not always hold; see, for example, Núñez *et al.* (1998). This is, in particular, the case for non-Abelian gauge fields. In addition to charges at spatial infinity, such 'coloured black holes' have to be characterized by additional variables, and it is necessary to approach the hole to

[1] Black holes could also have a magnetic-monopole charge, but this possibility will not be considered here.

determine them. The physical reason for the occurrence of such solutions is the non-linear character of these gauge fields. Fields in regions closer to the black hole (that would otherwise be swallowed by the hole) are tied to fields far away from the hole (that would otherwise be radiated away) to reach an equilibrium situation. In most examples this equilibrium is, however, unstable and the corresponding black-hole solution does not represent a physical solution. Since classical non-Abelian fields have never been observed (the description of objects such as quarks necessarily requires quantized gauge fields which, due to confinement, have no macroscopic limit), they will not be taken into account in the subsequent discussion.

In 1971, Stephen Hawking proved an important theorem about stationary black holes: their area can never decrease with time. More precisely, he showed that for a predictable black hole satisfying $R_{ab}k^ak^b \geq 0$ for all null k^a, the surface area of the *future* event horizon *never* decreases with time. A 'predictable' black hole is one for which the cosmic censorship hypothesis holds—this is thus a major assumption for the area law. Cosmic censorship assumes that all black holes occurring in nature have an event horizon, so that the singularity cannot be seen by far-away observers (the singularity is not 'naked'). The time asymmetry in this theorem comes into play because a statement is made about the future horizon, not the past horizon; the analogous statement for white holes would then be that the area of the past event horizon never increases. It is a feature of our universe that white holes seem to be absent, in contrast to black holes; cf. Section 10.2. It must be emphasized that the area law only holds in the classical theory, not in the quantum theory.

The area law exhibits a close analogy to the Second Law of Thermodynamics—there, the *entropy* can *never* decrease with time (for a closed system). However, the conceptual difference could not be more pronounced: while the Second Law is related to statistical behaviour, the area law is just a theorem in differential geometry.

Further support for this analogy is given by the existence of analogues to the other laws of thermodynamics. The Zeroth Law states that there exists a quantity, the temperature, that is constant over a body in thermal equilibrium. Does there exist an analogous quantity for a black hole? One can in fact prove that the surface gravity κ is constant over the event horizon (Wald 1984). For a Kerr black hole, κ is given by

$$\kappa = \frac{\sqrt{(GM)^2 - a^2}}{2GMr_+} \quad \xrightarrow{a \to 0} \quad \frac{1}{4GM} = \frac{GM}{R_\mathrm{S}^2}\,, \tag{7.1}$$

where

$$r_+ = GM + \sqrt{(GM)^2 - a^2}$$

denotes the location of the event horizon. In the Schwarzschild limit, one recognizes the well-known expression for the Newtonian gravitational acceleration. ($R_\mathrm{S} := 2GM$ in this case denotes the Schwarzschild radius.) One can show for a static black hole that κ is the limiting force that must be exerted at infinity to hold a unit test mass in place when it approaches the horizon. This justifies the name 'surface gravity'.

With a tentative formal relation between surface gravity and temperature, and between area and entropy, the question arises whether a First Law of Thermodynamics

Table 7.1 Analogies between the laws of thermodynamics and the laws of blackhole mechanics

Law	Thermodynamics	Stationary black holes
Zeroth	T constant over a body in thermal equilibrium	κ constant on the horizon of a black hole
First	$\mathrm{d}E = T\,\mathrm{d}S - p\,\mathrm{d}V + \mu\,\mathrm{d}N$	$\mathrm{d}M = \dfrac{\kappa}{8\pi G}\mathrm{d}A + \Omega_{\mathrm{H}}\,\mathrm{d}J + \Phi\,\mathrm{d}q$
Second	$\mathrm{d}S \geq 0$	$\mathrm{d}A \geq 0$
Third	$T = 0$ cannot be reached	$\kappa = 0$ cannot be reached

can be proved. This can in fact be achieved, and the result for a Kerr–Newman black hole is

$$\mathrm{d}M = \frac{\kappa}{2\pi}\frac{\mathrm{d}A}{4G} + \Omega_{\mathrm{H}}\,\mathrm{d}J + \Phi\,\mathrm{d}q\,, \tag{7.2}$$

where A, Ω_{H}, and Φ denote the area of the event horizon, the angular velocity of the black hole, and the electrostatic potential, respectively. This relation can be obtained by two conceptually different methods: a *physical-process version*, in which a stationary black hole is altered by infinitesimal physical processes, and an *equilibrium-state version*, in which the areas of two stationary black-hole solutions of Einstein's equations are compared. Both methods lead to the same result (7.2).

Since M is the energy of the black hole, (7.2) is the analogue of the First Law of Thermodynamics given by

$$\mathrm{d}E = T\,\mathrm{d}S - p\,\mathrm{d}V + \mu\,\mathrm{d}N. \tag{7.3}$$

'Modern' derivations of (7.2) make use of both Hamiltonian and Lagrangian methods of GR. For example, the First Law follows from an arbitrary diffeomorphism-invariant theory of gravity whose field equations can be derived from a Lagrangian; see Wald (2001) and the references therein.

What about the Third Law of Thermodynamics? A 'physical-process version' was proved by Israel (1986)—it is impossible to reach $\kappa = 0$ in a finite number of steps, although it is unclear whether this is true under all circumstances (Farrugia and Hajicek 1979). This corresponds to the 'Nernst version' of the Third Law. The stronger 'Planck version', which states that the entropy goes to zero (or a material-dependent constant) if the temperature approaches zero, does not seem to hold. The above analogies are summarized in Table 7.1.

7.1.2 Hawking and Unruh radiation

What is the meaning of black-hole temperature and entropy? According to classical GR, a black hole cannot radiate and, therefore, the temperature can only have a formal meaning. Important steps towards the interpretation of black-hole entropy were made by Bekenstein (1973); cf. also Bekenstein (2001). He argued that the Second

Law of Thermodynamics would only be valid if a black hole possessed an entropy $S_{\rm BH}$; otherwise, one could lower the entropy in the universe by just throwing matter possessing a certain amount of entropy into a black hole. Comparing (7.2) with (7.3), one recognizes that black-hole entropy must be a function of the area, $S_{\rm BH} = f(A)$. Since the temperature must be positive, one must demand that $f'(A) > 0$. The simplest case, $f(A) \propto \sqrt{A}$, that is, $S_{\rm BH} \propto M$, would violate the Second Law because if two black holes merged, the mass of the resulting hole would be smaller than the sum of the masses of the original holes (due to energy emission through gravitational waves). With some natural assumptions, one can conclude that $S_{\rm BH} \propto A/l_{\rm P}^2$ (Bekenstein 1973, 2001). Note that Planck's constant $\hbar$ has entered the scene through the Planck length. This has happened because no fundamental length scale can be constructed from G and c alone. A sensible interpretation of black-hole temperature and entropy thus cannot be obtained in pure GR—quantum theory has to be taken into account.

Thus, one can write

$$T_{\rm BH} \propto \frac{\hbar\kappa}{k_{\rm B}}, \qquad S_{\rm BH} \propto \frac{k_{\rm B}A}{G\hbar}, \tag{7.4}$$

and the important question is how the proportionality factor can be determined. This was achieved in the important paper by Hawking (1975). The key ingredient in Hawking's discussion is the behaviour of *quantum* fields on the background of an object collapsing to form a black hole. Like the situation with an external electric field (the Schwinger effect), there is no uniquely defined notion of a *vacuum*. This leads to the occurrence of particle creation. The peculiarity of the black-hole case is the *thermal* distribution of the particles created, which is due to the presence of an event horizon.

It is helpful to understand the 'Hawking effect' by first considering an analogous effect in flat space–time: a uniformly accelerated observer experiences the Minkowski vacuum as being filled with thermal particles; cf. Unruh (1976). Whereas all inertial observers in Minkowski space agree on the notion of a vacuum (and therefore on particles), this is no longer true for *non-inertial* observers.

Let us consider an observer who is uniformly accelerating along the X-direction in (1+1)-dimensional Minkowski space–time (Fig. 7.1). The Minkowski cartesian coordinates are labelled here by uppercase letters. The orbit of this observer is the hyperbola shown in the figure. One recognizes that, as in the Kruskal diagram for the Schwarzschild metric, the observer encounters a horizon (here called the 'acceleration horizon'). There is, however, no singularity behind this horizon. Region I is a globally hyperbolic space–time on its own—called the *Rindler space–time*. This space–time can be described by coordinates (τ, ρ), which are connected to the cartesian coordinates via the transformation

$$\begin{pmatrix} T \\ X \end{pmatrix} = \rho \begin{pmatrix} \sinh a\tau \\ \cosh a\tau \end{pmatrix}, \tag{7.5}$$

where a is a constant (the orbit in Fig. 7.1 describes an observer with acceleration a, who has $\rho = 1/a$). Since

$$\mathrm{d}s^2 = \mathrm{d}T^2 - \mathrm{d}X^2 = a^2\rho^2\,\mathrm{d}\tau^2 - \mathrm{d}\rho^2, \tag{7.6}$$

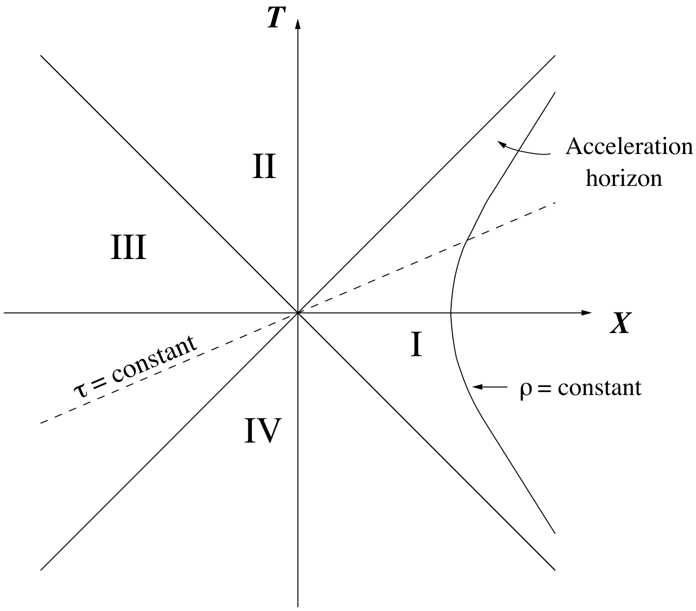

Fig. 7.1 Uniformly accelerated observer in Minkowski space.

the orbits $\rho =$ constant are also orbits of a time-like Killing field $\partial/\partial\tau$. It is clear that τ corresponds to the external Schwarzschild coordinate t and that ρ corresponds to r. As in the Kruskal case, $\partial/\partial\tau$ becomes space-like in regions II and IV.

The analogy with the Kruskal case becomes even more transparent if the Schwarzschild metric is expanded around the horizon at $r = 2GM$. Introducing there a new coordinate ρ via $\rho^2/(8GM) = r - 2GM$ and recalling (7.1), one has

$$\mathrm{d}s^2 \approx \kappa^2\rho^2\,\mathrm{d}t^2 - \mathrm{d}\rho^2 - \frac{1}{4\kappa^2}\mathrm{d}\Omega^2. \tag{7.7}$$

Comparison with (7.6) shows that the first two terms on the right-hand side of (7.7) correspond exactly to the Rindler space–time (7.6) with the acceleration a replaced by the surface gravity κ. The last term[2] in (7.7) describes a two-sphere with radius $(2\kappa)^{-1}$.

How does the accelerating observer experience the standard Minkowski vacuum $|0\rangle_{\mathrm{M}}$? The key point is that the vacuum is a *global* state with correlations between regions I and III in Fig. 7.1 (similar to Einstein–Podolsky–Rosen correlations), but that the accelerated observer is restricted to region I. Considering for simplicity the case of a massless scalar field, the global vacuum state comprising regions I and III can be written in the form

$$|0\rangle_{\mathrm{M}} = \prod_{\omega}\sqrt{1-\mathrm{e}^{-2\pi\omega a^{-1}}}\sum_{n}\mathrm{e}^{-n\pi\omega a^{-1}}|n_\omega^{\mathrm{I}}\rangle\otimes|n_\omega^{\mathrm{III}}\rangle, \tag{7.8}$$

[2]It is this term that is responsible for the non-vanishing curvature of (7.7) compared with the flat-space metric (7.6), whose extension into the (neglected) other dimensions would be just $-\mathrm{d}Y^2 - \mathrm{d}Z^2$.

where $|n_\omega^{\mathrm{I}}\rangle$ and $|n_\omega^{\mathrm{III}}\rangle$ are n-particle states with frequency $\omega = |\mathbf{k}|$ in regions I and III, respectively. These n-particle states are defined with respect to the 'Rindler vacuum', which is the vacuum defined by an accelerating observer. The expression (7.8) is an example of the Schmidt decomposition of two entangled quantum systems; see, for example, Joos *et al.* (2003). Note also the analogy of (7.8) to a BCS state in the theory of superconductivity; see, for example, Tinkham (2004).

For an observer restricted to region I, the state (7.8) cannot be distinguished, by operators with support in I only, from a density matrix that is found from (7.8) by tracing out all degrees of freedom in region III,

$$\begin{aligned}\rho_{\mathrm{I}} &:= \mathrm{tr}_{\mathrm{III}}|0\rangle_{\mathrm{M}}\langle 0|_{\mathrm{M}} \\ &= \prod_\omega \left(1 - \mathrm{e}^{-2\pi\omega a^{-1}}\right) \sum_n \mathrm{e}^{-2\pi n\omega a^{-1}} |n_\omega^{\mathrm{I}}\rangle\langle n_\omega^{\mathrm{I}}| . \end{aligned} \tag{7.9}$$

Note that the density matrix ρ_{I} has exactly the form corresponding to a thermal canonical ensemble with the Davies–Unruh temperature

$$T_{\mathrm{DU}} = \frac{\hbar a}{2\pi k_{\mathrm{B}}} \quad \approx 4.05 \times 10^{-23}\, a \left[\frac{\mathrm{cm}}{\mathrm{s}^2}\right] \mathrm{K}\ ; \tag{7.10}$$

cf. (1.36). An observer who is accelerating uniformly through Minkowski space thus sees a *thermal* distribution of particles. This is an important manifestation of the non-uniqueness of the vacuum state in quantum field theory, even for flat space–time. A more detailed discussion invoking models of particle detectors confirms this result (the 'Unruh effect'). A comprehensive review of the Unruh effect has been given in Crispino *et al.* (2008). Laser experiments designed to observe this effect are in preparation (Section 1.1.5).

We shall now turn to black holes. From the form of the line element near the horizon (7.7), one can already anticipate that—according to the equivalence principle—a black hole radiates with a temperature as specified in (7.10) with a replaced by κ. This is, in fact, what Hawking (1975) found. In his approach, he considered the situation depicted in Fig. 7.2: the vacuum modes of a quantum field are calculated on the background of a star collapsing to form a black hole.

Due to the dynamical background, an initial vacuum does not stay a vacuum but becomes a thermal state with respect to late-time observers. The temperature of this state is called the 'Hawking temperature' and is equal to

$$T_{\mathrm{BH}} = \frac{\hbar\kappa}{2\pi k_{\mathrm{B}}}\ ; \tag{7.11}$$

cf. (1.34). In the Heisenberg picture, this non-invariance of the vacuum is called 'particle creation' (more properly, 'field excitation'). In the Schrödinger picture, it corresponds to the process of squeezing of quantum states; cf. Grishchuk and Sidorov (1990) and Kiefer (2001*b*). More precisely, it is a two-mode squeezed state, which in this context is also called the 'Unruh vacuum'. It is a quantum state that exhibits entanglement between the inside and the outside of the black hole. As with (7.9), tracing out the interior part leads to a thermal spectrum with the temperature (7.11). It is

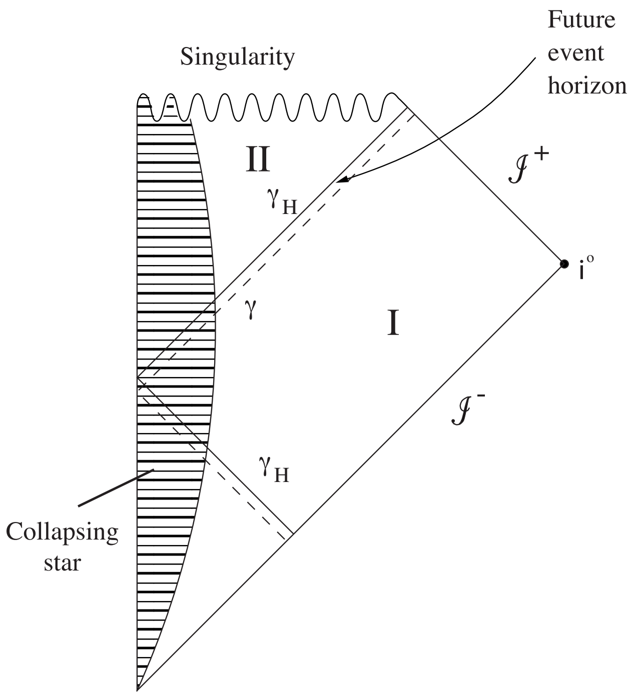

Fig. 7.2 Penrose diagram showing the collapse of a spherically symmetric star to form a black hole; γ denotes a light ray propagating from $\mathcal{J}^-$ through the collapsing star to $\mathcal{J}^+$. The limiting case is the ray $\gamma_{\rm H}$ which stays on the horizon.

true in general that tracing out one mode from a two-mode squeezed state yields a thermal density matrix; see, for example, Section 5.2.5 in Walls and Milburn (1994). The fact that the excited field modes are thermally distributed here follows from the presence of a horizon.

The temperature (7.11) refers to an observer at an infinite distance from the hole. For a finite distance, one has to modify this expression by a redshift factor. If the observer is static, he/she is in a state of acceleration, and therefore the temperature contains both the Hawking and the Unruh effect. Close to the horizon, only the Unruh effect remains, and a freely falling observer does not experience any temperature at all.

According to Hawking (1975), the number of particles emitted per second with total energy between E and $E + \mathrm{d}E$ is, for a Kerr–Newman black hole, given by

$$\mathrm{d}\dot{N}_s = \sum_{l,m} \frac{\mathrm{d}E}{2\pi\hbar} \frac{\Gamma_{slm}}{\mathrm{e}^{(E-m\hbar\Omega-q\Phi)/k_\mathrm{B}T_\mathrm{BH}} - (-1)^{2s}}, \tag{7.12}$$

where s denotes the particle spin, l is the orbital angular momentum of the particle, and $m\hbar$ is the axial quantum number; l and m arise from the spherical harmonics Y_{lm}.

The terms Γ_{slm} are the *greybody factors*. These are dimensionless numbers that give the absorption probability for the emitted particle species. They take into account the fact that some of the particle modes are back-scattered into the black hole by the space–time curvature. The corresponding part therefore does not reach infinity and, consequently, leads to a modification of the black-body spectrum in (7.12).

In the following, we shall assume the special case of the Schwarzschild metric, where the surface gravity reads (reinserting c) $\kappa = c^4/4GM$. In this case, (7.11) becomes the expression (1.35), which can also be written as

$$k_\mathrm{B}T_\mathrm{BH} = \frac{\hbar c^3}{8\pi GM} \approx 1.06 \left(\frac{M}{10^{10}\ \mathrm{g}}\right)^{-1}\ \mathrm{TeV}. \tag{7.13}$$

Massless elementary particles are emitted for all temperatures. The known cases are $s = 1$ (photon) and $s = 2$ (graviton). One can also include the approximately massless neutrinos ($s = 1/2$). Massive particles are emitted in significant numbers once the peak in the energy distribution (7.12) becomes comparable to the particle rest mass. Because of the tail of the distribution (7.12), a non-zero contribution is, however, always present. In this way, the distribution (7.12) can be connected with the particle physics at the corresponding energy scale.

From (7.12), one obtains for the total luminosity of a Schwarzschild black hole the expression

$$L := -\frac{\mathrm{d}M}{\mathrm{d}t} = -\sum_{l,s} \mathrm{d}E\ \frac{E}{c^2} \frac{\Gamma_{sl}}{\mathrm{e}^{E/k_\mathrm{B}T_\mathrm{BH}} - (-1)^{2s}}, \tag{7.14}$$

where the inserted factor E/c^2 is the mass of the emitted particle, and the sum over s is a sum over all species.[3]

The greybody factors are, in general, complicated functions of the mass and energy that can only be evaluated numerically. In special limiting cases, however, one can also obtain simple analytic expressions (Page 1976, MacGibbon and Webber 1990). The limit of high energy ($E \gg k_\mathrm{B}T_\mathrm{BH}$) corresponds to the geometric-optics limit, which holds universally for all species. One finds in this limit

$$\Gamma_s \approx \frac{27G^2M^2E^2}{\hbar^2c^6} \approx 3.82 \times 10^{-2} \left(\frac{M}{10^{10}\ \mathrm{g}}\right)^2 \left(\frac{E}{\mathrm{TeV}}\right)^2, \tag{7.15}$$

where Γ_s is generally defined by

$$\Gamma_s := \sum_l \Gamma_{sl} = \frac{E^2\sigma_s(M,E)}{\pi\hbar^2c^2}, \tag{7.16}$$

and σ_s denotes the absorption cross-section.

[3] Sometimes the factor $2l+1$ arising from the sum over the axial quantum number is extracted from Γ_{sl}.

In the limit of low energy ($E \gg k_B T_{BH}$), different results are obtained for different species. For a massless $s = 0$ particle one finds, for example,

$$\Gamma_0 \approx \frac{16G^2M^2E^2}{\hbar^2 c^6} \approx 2.26 \times 10^{-2} \left(\frac{M}{10^{10}\ \mathrm{g}}\right)^2 \left(\frac{E}{\mathrm{TeV}}\right)^2 \ ;$$

the peak in the particle flux at infinity is found for $s = 0$ at $E \approx 2.81 k_B T_{BH}$.

One can estimate the lifetime of a black hole by making the plausible assumption that the decrease in mass is equal to the energy radiated to infinity. This corresponds to a heuristic implementation of the back reaction of the Hawking radiation on the black hole. (Without this back reaction, energy would not be conserved.) Using the Stefan–Boltzmann law, one gets

$$\frac{\mathrm{d}M}{\mathrm{d}t} \propto -AT_{BH}^4 \propto -M^2 \times \left(\frac{1}{M}\right)^4 = -\frac{1}{M^2},$$

which when integrated yields

$$t(M) \propto (M_0^3 - M^3) \approx M_0^3. \tag{7.17}$$

Here M_0 is the initial mass. It has been assumed that after the evaporation, $M \ll M_0$. Very roughly, the lifetime of a black hole is thus given by

$$\tau_{BH} \approx \left(\frac{M_0}{m_P}\right)^3 t_P \approx 10^{65} \left(\frac{M_0}{M_\odot}\right)^3 \text{years}. \tag{7.18}$$

A more precise value for the lifetime of a Schwarzschild black hole can be obtained by integrating (7.14) for the Standard Model of particle physics (MacGibbon 1991). One finds

$$\tau_{BH} \approx 407 \left(\frac{f(M_0)}{15.35}\right)^{-1} \left(\frac{M}{10^{10}\ \mathrm{g}}\right)^3 \ \mathrm{s} \approx 6.24 \times 10^{-27} M_0^3[\mathrm{g}] f^{-1}(M_0)\ \mathrm{s}, \tag{7.19}$$

where $f(M_0)$ is a measure of the number of emitted particle species (e.g. $f = 0.267$ for a massless $s = 0$ particle); it is normalized to $f(M_0) = 1$ for $M_0 \gg 10^{17}$ g when only (effectively) massless particles are emitted. If one sums over the contributions from all particles in the Standard Model up to an energy of about 1 TeV, one has $f(M) = 15.35$, which motivates the occurrence of this number in (7.19) (Carr *et al.* 2010).

The evaporation situation becomes complicated when the black hole is hot enough to emit hadrons. This happens for $T_{BH} \geq \Lambda_{QCD} \approx 250$–$300$ MeV, where Λ_{QCD} is the QCD confinement scale. The black hole then emits quarks and gluon jets that condense into the observed hadrons at some distance from the black hole. The details of these QCD processes are dealt with in MacGibbon and Webber (1990) and MacGibbon (1991).

It is interesting to know the initial mass of a black hole formed in the very early Universe (see Section 7.7) that will evaporate today, that is, for which the lifetime

(7.19) coincides with the age of our Universe. Using the value 13.7×10^9 years for this age, one finds from (7.19) that this happens for a mass

$$M_0 \approx 5 \times 10^{14} \text{ g} =: M_*. \tag{7.20}$$

This is important for the possible detection of primordial black holes (Section 7.7). The mass M_* corresponds to the mass of a small asteroid.[4] From (7.13), one finds that this corresponds to a Hawking temperature of about 20 MeV.

The Schwarzschild radius of a black hole with mass M and Hawking temperature $T_{\rm BH}$ can be expressed as

$$R_{\rm S} \approx 1.48 \times 10^{-18} \left(\frac{M}{10^{10} \text{ g}}\right) \text{ cm} \approx 1.57 \times 10^{-18} \left(\frac{k_{\rm B}T_{\rm BH}}{\text{TeV}}\right)^{-1} \text{ cm}.$$

A black hole with mass M_* thus has a Schwarzschild radius of order 10^{-13} cm, which corresponds to the size of an atomic nucleus.

MacGibbon *et al.* (2008) also considered the reduced de Broglie wavelength for directly emitted particles of energy E,

$$\lambda = \frac{\hbar c}{E} \approx 1.97 \times 10^{-17} \left(\frac{E}{\text{TeV}}\right)^{-1} \text{ cm}.$$

Since the peak energy of an emitted particle is of order 3–5 $k_{\rm B}T_{\rm BH}$, this wavelength is larger than the black hole. For example, for the case of a massless $s = 0$ particle with peak energy $E \approx 2.81 k_{\rm B}T_{\rm BH}$ considered above, one finds $\lambda \approx 4.5 R_{\rm S}$. MacGibbon *et al.* (2008) also emphasized that the particles can be considered to have 'evaporated' at a distance $\approx 3R_{\rm S}/2$ from the black hole, since the outgoing solution of the wave equation then resembles the solution at infinity.

Hawking used a semiclassical approximation in which the non-gravitational fields are quantum but the gravitational field is treated as an external classical field; cf. Section 5.4. This approximation is expected to break down when the black-hole mass approaches the Planck mass, that is, after a time given by (7.18). Only a full theory of quantum gravity can describe the final stages of black-hole evaporation. This would necessarily include the full back-reaction effect of the Hawking radiation on the black hole.

The original derivation of the Hawking effect deals with non-local 'particle' modes (Hawking 1975). Alternatively, one can base the analysis solely on the local behaviour of correlation functions by inspecting their time development from the past into the future (Fredenhagen and Haag 1990).

As an intermediate step towards full quantum gravity, one might consider the heuristic 'semiclassical' Einstein equations discussed in Section 1.2; see (1.37). This enables one to take into account back-reaction effects at the semiclassical level. The evaluation of $\langle T_{ab}\rangle$—which requires regularization and renormalization—is a difficult subject on its own (Frolov and Novikov 1998). The renormalized value for $\langle T_{ab}\rangle$ is

[4] The minor planet Apophis, for example, which will come rather close to the Earth in 2029, has a mass of about 2.7×10^{13} g.

essentially unique (its ambiguities can be absorbed into coupling constants) if certain sensible requirements are imposed; cf. Section 2.2.4. By evaluating the components of the renormalized $\langle T_{ab}\rangle$ near the horizon, one finds that there is a flux of *negative energy* into the hole. This follows from the Unruh vacuum described above. Clearly, this leads to a decrease of the mass. These negative energies represent a typical quantum effect and are well known from the accurately measured Casimir effect. This occurrence of negative energies is also responsible for the breakdown of the classical area law in quantum theory.

The negative flux near the horizon also lies at the heart of the 'tunnelling' interpretation of Hawking radiation (Parikh and Wilczek 2000). In the vacuum, virtual pairs of 'particles' are created and destroyed. However, close to the horizon, one partner of this virtual pair might fall (tunnel) into the black hole, thereby liberating the other partner to become a real particle and to escape to infinity as Hawking radiation. The global quantum field exhibits quantum entanglement between the inside and outside of the black hole, similarly to the case of the accelerated observer discussed above. A detailed account of the tunnelling interpretation is given in Vanzo *et al.* (2011).

In the case of an eternal Schwarzschild black hole, where both past and future horizons exist, there exists a distinguished quantum state that describes the equilibrium of the black hole with thermal radiation at the Hawking temperature. This state is also called the 'Hartle–Hawking vacuum' (Hartle and Hawking 1976). It is directly analogous to the Minkowski vacuum in (7.8).

One can also give explicit expressions for the Hawking temperature (7.11) in the case of rotating and charged black holes. For the Kerr solution, one has

$$k_{\mathrm{B}}T_{\mathrm{BH}} = \frac{\hbar\kappa}{2\pi} = 2\left(1+\frac{GM}{\sqrt{G^2M^2-a^2}}\right)^{-1}\frac{\hbar}{8\pi GM} < \frac{\hbar}{8\pi GM}. \tag{7.21}$$

Rotation thus reduces the Hawking temperature. For the Reissner–Nordström solution (describing a charged spherically symmetric black hole), one has

$$k_{\mathrm{B}}T_{\mathrm{BH}} = \frac{\hbar}{8\pi GM}\left(1-\frac{G^2q^4}{r_+^4}\right) < \frac{\hbar}{8\pi GM}. \tag{7.22}$$

Electric charge thus also reduces the Hawking temperature. For an extremal black hole, $r_+ = GM = \sqrt{G}|q|$, and therefore $T_{\mathrm{BH}} = 0$.

A conceptual problem in the traditional derivation of the Hawking effect should be mentioned. As treated, the propagation of the modes through the collapse phase involves modes with very high frequency; in fact, the corresponding wavelengths can be much smaller than the Planck length, and one might wonder whether one would not need the full quantum theory of gravity for the calculation of the Hawking effect. This is called the 'trans-Planckian problem'; cf. Jacobson (2003) and the references therein. An analogous version of this problem also occurs in inflationary cosmology. There is no general consensus on the relevance of this problem in the literature; however, since the Hawking effect can be derived in various different ways, not necessarily involving the propagation of modes through the collapse phase,[5] it seems that the trans-Planckian problem does not pose a threat to the Hawking effect.

[5]One example is the work by Fredenhagen and Haag (1990) mentioned above.

To discuss problems like these, it may be helpful to mention that there exist interesting analogues to black-hole quantum effects in condensed-matter physics; see Barceló *et al.* (2011) for a detailed review. The study of these analogues was motivated partly by the trans-Planckian problem, because an understanding of a similar feature in analogous models could shed some light on this problem. As far as the Hawking effect is concerned, an important property that these models should have is the presence of an apparent horizon. Such analogous models might be provided by superfluid helium or Bose–Einstein condensates (which seem to provide the most promising model). Typically, these horizons are *acoustic* horizons: a horizon is defined in this case to be a two-dimensional surface on which the normal component of the fluid velocity is everywhere equal to the local speed of *sound*.

7.1.3 The Bekenstein–Hawking entropy

After the Hawking temperature has been calculated, the entropy is also given. From the First Law (7.2), one finds the 'Bekenstein–Hawking entropy'

$$S_{\rm BH} = \frac{k_{\rm B} A}{4G\hbar}\,, \tag{7.23}$$

in which the unknown factor in (7.4) has now been fixed. For the special case of a Schwarzschild black hole, this yields

$$S_{\rm BH} = \frac{k_{\rm B}\pi R_{\rm S}^2}{G\hbar} \approx 1.07 \times 10^{77} k_{\rm B} \left(\frac{M}{M_\odot}\right)^2 . \tag{7.24}$$

It can be easily estimated that $S_{\rm BH}$ is much greater than the entropy of the star that collapsed to form the black hole. The entropy of the Sun, for example, is given approximately by $S_\odot \approx 10^{57} k_{\rm B}$, whereas the entropy of a solar-mass black hole is about $10^{77} k_{\rm B}$, which is 20 orders of magnitude larger (recall that the number of states is given by $\exp(S/k_{\rm B})$).

Can a physical interpretation of this huge discrepancy be given? In the above discussion, the laws of black-hole mechanics have been treated as phenomenological thermodynamical laws. The central open question therefore is: can $S_{\rm BH}$ be derived from quantum-statistical considerations? This would mean that $S_{\rm BH}$ could be calculated from a Gibbs-type formula as a statistical-mechanical entropy $S_{\rm SM}$ according to

$$S_{\rm BH} \stackrel{?}{=} -k_{\rm B}{\rm tr}(\rho \ln \rho) =: S_{\rm SM}\,, \tag{7.25}$$

where ρ denotes an appropriate density matrix; $S_{\rm BH}$ would then somehow correspond to the number of quantum microstates that are consistent with the macrostate of the black hole. According to the no-hair theorem, the macrostate is uniquely characterized by its mass, angular momentum, and charge. A symbolic picture inspired by John Wheeler is shown in Figure 7.3: the (unknown) microstates are symbolized by a distribution of bits on the event horizon of a black hole.

Some important questions for the interpretation of the entropy are:

- Does $S_{\rm BH}$ correspond to states hidden behind the horizon?

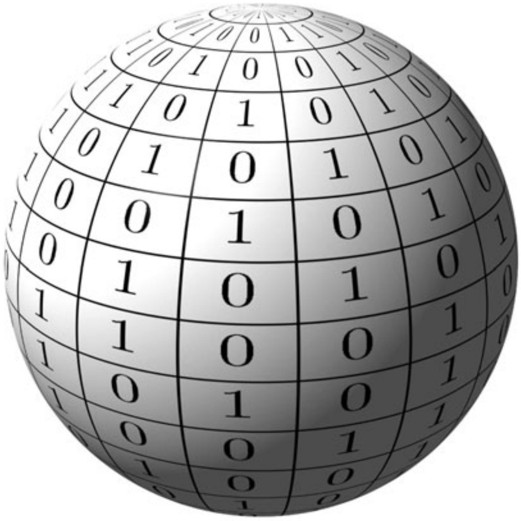

Fig. 7.3 Attachment of bits to the event horizon of a black hole, symbolizing the counting of microstates. Adapted from Wheeler (1990).

- Or does S_{BH} correspond to the number of possible initial states from which the black hole might have formed?
- What are the microscopic degrees of freedom?
- Where are they located (if at all)?
- Can one understand the universality of the result?
- What happens to S_{BH} after the black hole has evaporated?
- Is the entropy a 'one-loop' or a 'tree-level' effect?

The attempts to calculate S_{BH} by state counting are often done in the 'one-loop limit' of quantum field theory in curved space–time—this is the limit where gravity is classical but non-gravitational fields are fully quantum; cf. Sections 2.2 and 5.4. Equation (7.11) has been derived in this limit. Equation (7.23) can be calculated from the so-called 'tree-level' approximation of the theory, where only the gravitational degrees of freedom are taken into account. Usually, a saddle-point approximation to a Euclidean path integral is calculated. Such derivations are, however, equivalent to derivations within classical thermodynamics; cf. Wald (2001).

The emergence of the thermal nature of black-hole radiation has led to the discussion of the *information-loss problem*; cf. Section 7.6. If the black hole were to evaporate completely and leave only thermal radiation behind, one would have a conflict with established principles in quantum theory: any initial state (in particular, a pure state) would evolve into a mixed state. In ordinary quantum theory, this is forbidden by the unitary evolution of the total closed system. A theory of quantum gravity should give a definite answer to the question of whether unitarity (with respect to an outside observer) is preserved or not.

In the following subsections, as well as in Chapter 9, we shall discuss attempts to describe black holes and their evolution within quantum gravity.

7.2 Canonical quantization of the Schwarzschild black hole

What can the methods of Chapters 5 and 6 say about the quantum behaviour of Schwarzschild black holes? In the following, we shall present an outline of the canonical quantization procedure for spherically symmetric systems. This was developed in the connection representation (using the complex version of Ashtekar's variables) by Thiemann and Kastrup (1993) in the spirit of the Dirac approach discussed in Section 5.2.2. In the reduced version described in Section 5.2.1, it was developed by Kastrup and Thiemann (1994) using the connection variables and by Kuchař (1994) using the geometrodynamical representation. An extension of the latter work to charged black holes (Reissner–Nordström solutions) can be found in Louko and Winters-Hilt (1996). In this subsection, we discuss the geometrodynamical approach and follow, with elaborations, the presentation in Kiefer (1998).

Models of this kind are often called *midisuperspace models* (see e.g. Barbero and Villaseñor 2010) because the complexity of their configuration space lies between full superspace (Section 4.2.5) and the finite-dimensional models ('minisuperspace models') discussed in quantum cosmology (Chapter 8).

7.2.1 Classical formalism

The starting point is the ansatz for a general spherically symmetric metric on $\mathbb{R}\times\mathbb{R}\times S^2$,

$$\mathrm{d}s^2 = -N^2(r,t)\,\mathrm{d}t^2 + L^2(r,t)(\mathrm{d}r + N^r(r,t)\,\mathrm{d}t)^2 + R^2(r,t)\,\mathrm{d}\Omega^2. \tag{7.26}$$

The lapse function N encodes the possibility of performing arbitrary reparametrizations of the time parameter, while the shift function N^r is responsible for reparametrizations of the radial coordinate (this is the only freedom in performing spatial coordinate transformations that is left after spherical symmetry is imposed). The parameter r is merely a label for the spatial hypersurfaces; if the hypersurface extends from the left to the right wedge in the Kruskal diagram, one takes $r\in(-\infty,\infty)$. If the hypersurface originates at the bifurcation point, where the past and future horizons meet, one has $r\in(0,\infty)$. If one has in addition a spherically symmetric electromagnetic field, one makes the following ansatz for the one-form potential:

$$A = \phi(r,t)\,\mathrm{d}t + \Gamma(r,t)\,\mathrm{d}r. \tag{7.27}$$

In the Hamiltonian formulation, ϕ, N, and N^r are Lagrange multipliers whose variations yield the constraints of the theory. Variation of the Einstein–Hilbert action with respect to N yields the Hamiltonian constraint, which for the spherically symmetric model is given by

$$\mathcal{H}_\perp = \frac{G}{2}\frac{LP_L^2}{R^2} - G\frac{P_LP_R}{R} + \frac{LP_\Gamma^2}{2R^2} + G^{-1}V^{\mathrm{g}} \approx 0\ , \tag{7.28}$$

where the gravitational potential term reads

$$V^{\mathrm{g}} = \frac{RR''}{L} - \frac{RR'L'}{L^2} + \frac{R'^2}{2L} - \frac{L}{2}. \tag{7.29}$$

(A prime denotes differentiation with respect to r.) Variation with respect to N^r yields one (radial) diffeomorphism constraint,

$$\mathcal{H}_r = P_R R' - L P_L' \approx 0. \tag{7.30}$$

One recognizes from this constraint that R transforms as a scalar, while L transforms as a scalar density.

Variation of the action with respect to ϕ yields as usual the Gauss constraint

$$\mathcal{G} = P_\Gamma' \approx 0. \tag{7.31}$$

The constraint (7.30) generates radial diffeomorphisms for the fields R, L, and their canonical momenta. It does not generate diffeomorphisms for the electromagnetic variables. This can be taken into account if one uses the multiplier $\tilde{\phi} = \phi - N^r\Gamma$ instead of ϕ and varies with respect to $\tilde{\phi}$ (Louko and Winters-Hilt 1996), but for the present purpose it is sufficient to stick to the above form (7.30).

The model of spherically symmetric gravity can be embedded into a whole class of models usually referred to as 'two-dimensional dilaton gravity theories'. This terminology comes from effective two-dimensional theories (usually motivated by string theory), which contain in the gravitational sector a scalar field (the 'dilaton') in addition to the two-dimensional metric (of which only the conformal factor is relevant). An example is the 'CGHS model' defined in (5.60), within which one can address the issues of Hawking radiation and back reaction. This model is classically soluble even if a further, conformally coupled, scalar field is included. The canonical formulation of this model can be found, for example, in Louis-Martinez *et al.* (1994) and Demers and Kiefer (1996). The dilaton field is analogous to the field R above, while the conformal factor of the two-dimensional metric is analogous to L.

Let us consider now the boundary conditions for $r \to \infty$. One has in particular

$$L(r,t) \to 1 + \frac{GM(t)}{r}, \ R(r,t) \to r, \ N \to N(t), \tag{7.32}$$

as well as

$$P_\Gamma(r,t) \to q(t), \quad \phi(r,t) \to \phi(t). \tag{7.33}$$

From the variation with respect to L, one then finds the boundary term $\int \mathrm{d}t\ N\delta M$. In order to avoid the unwanted conclusion $N = 0$ (no evolution at infinity), one has to compensate this term in advance by adding the boundary term

$$-\int \mathrm{d}t\ NM$$

to the classical action. Note that M is just the ADM mass. The need to include such a boundary term was recognized by Regge and Teitelboim (1974); cf. Section 4.2.4. Similarly, for charged black holes one has to add the term

$$-\int \mathrm{d}t\ \phi q$$

to compensate for $\int \mathrm{d}t\ \phi\delta q$, which arises from varying P_Γ. If one wished instead to consider q as a given external parameter, this boundary term would be unnecessary.

As long as restriction is made to eternal black holes, appropriate canonical transformations allow one to simplify the classical constraint equations considerably (Kuchař 1994; Louko and Winters-Hilt 1996). One gets

$$(L, P_L; R, P_R; \Gamma, P_\Gamma) \longrightarrow (\mathcal{M}, P_{\mathcal{M}}; \mathcal{R}, P_{\mathcal{R}}; Q, P_Q)\,.$$

In particular,

$$\mathcal{M}(r,t) = \frac{P_\Gamma^2 + P_L^2}{2R} + \frac{R}{2}\left(1 - \frac{R'^2}{L^2}\right) \stackrel{r\to\infty}{\longrightarrow} M(t), \tag{7.34}$$

$$Q(r,t) = P_\Gamma \stackrel{r\to\infty}{\longrightarrow} q(t) \tag{7.35}$$

($\mathcal{R} = R$; the expression for $P_{\mathcal{R}}$ is somewhat lengthy and will not be given here.)

The new constraints, which are equivalent to the old ones, read

$$\mathcal{M}' = 0 \quad \Rightarrow \quad \mathcal{M}(r,t) = M(t), \tag{7.36}$$

$$Q' = 0 \quad \Rightarrow \quad Q(r,t) = q(t), \tag{7.37}$$

$$P_{\mathcal{R}} = 0\ . \tag{7.38}$$

Note that $N(t)$ and $\phi(t)$ are prescribed functions that must not be varied; otherwise one would get the unwanted restriction $M = 0 = q$. A variation is allowed if the action is parametrized, bringing in new dynamical variables,

$$\begin{aligned} N(t) &=: \dot{\tau}(t), \\ \phi(t) &=: \dot{\lambda}(t). \end{aligned} \tag{7.39}$$

Here, τ is the proper time that is measured with standard clocks at infinity, and λ is the variable conjugate to the charge; λ is therefore connected with the electromagnetic gauge parameter at the boundaries. In the canonical formalism one has to introduce momenta conjugate to these variables, which will be denoted by π_τ and π_λ, respectively. This, in turn, requires the introduction of additional constraints linear in the momenta,

$$\mathcal{C}_\tau = \pi_\tau + M \approx 0, \tag{7.40}$$

$$\mathcal{C}_\lambda = \pi_\lambda + q \approx 0, \tag{7.41}$$

which have to be added to the action:

$$-\int \mathrm{d}t\ M\dot{\tau} \quad \to \quad \int \mathrm{d}t\ (\pi_\tau\dot{\tau} - N\mathcal{C}_\tau), \tag{7.42}$$

$$-\int \mathrm{d}t\ q\dot{\lambda} \quad \to \quad \int \mathrm{d}t\ (\pi_\lambda\dot{\lambda} - \phi\mathcal{C}_\lambda). \tag{7.43}$$

The remaining constraints in this model are thus (7.38), and (7.40) and (7.41).

7.2.2 Quantization

One then quantizes as in Chapter 5 by acting with an operator version of the constraints on wave functionals $\Psi[\mathcal{R}(r);\tau,\lambda)$. Since (7.38) leads to $\delta\Psi/\delta\mathcal{R}=0$, one is left with a purely quantum-*mechanical* wave function $\psi(\tau,\lambda)$. One could call this a 'quantum Birkhoff theorem'. The implementation of the constraints (7.40) and (7.41) then yields

$$\frac{\hbar}{\mathrm{i}}\frac{\partial\psi}{\partial\tau}+M\psi=0, \tag{7.44}$$

$$\frac{\hbar}{\mathrm{i}}\frac{\partial\psi}{\partial\lambda}+q\psi=0, \tag{7.45}$$

which can be readily solved to give

$$\psi(\tau,\lambda)=\chi(M,q)\mathrm{e}^{-\mathrm{i}(M\tau+q\lambda)/\hbar} \tag{7.46}$$

with an arbitrary function $\chi(M,q)$. Note that M and q are kept fixed here because, up to now, we have restricted our attention to a single semiclassical component of the wave function (eigenstates of mass and charge).

If the hypersurface goes through the whole Kruskal diagram of the eternal black hole, only the boundary term at $r\to\infty$ (and an analogous one for $r\to-\infty$) contributes. Of particular interest in the black-hole case, however, is the situation when the surface originates at the 'bifurcation surface' ($r\to 0$) of the past and future horizons. This makes sense, since data on such a surface suffice to construct the whole right Kruskal wedge, which is all that is accessible to an observer in this region. Moreover, this mimics the situation in which a black hole is formed by collapse and the regions III and IV of the Kruskal diagram are absent.

What are the boundary conditions to be adopted at $r\to 0$? They are chosen in such a way that the classical solutions have a non-degenerate horizon and the hypersurfaces start at $r=0$ asymptotically to the hypersurfaces of constant Killing time (Louko and Whiting 1995). In particular,

$$N(r,t)=N_1(t)r+\mathcal{O}(r^3), \tag{7.47}$$

$$L(r,t)=L_0(t)+\mathcal{O}(r^2), \tag{7.48}$$

$$R(r,t)=R_0(t)+R_2(t)r^2+\mathcal{O}(r^4). \tag{7.49}$$

As with the situation at $r\to\infty$, variation leads to a boundary term at $r=0$,

$$-N_1R_0(GL_0)^{-1}\delta R_0\ .$$

If $N_1\neq 0$, this term must be subtracted ($N_1=0$ corresponds to the case of extremal holes, $|q|=\sqrt{G}M$, which is characterized by $\partial N/\partial r(r=0)=0$). Introducing the notation $N_0:=N_1/L_0$, the boundary term to be added to the classical action is

$$(2G)^{-1}\int \mathrm{d}t\ N_0R_0^2\ .$$

The quantity

$$\alpha := \int_{t_1}^{t} \mathrm{d}t \; N_0(t) \tag{7.50}$$

can be interpreted as a 'rapidity' because it boosts n^a, the normal vector to the $t =$ constant hypersurfaces, in accordance with

$$n^a(t_1)n_a(t) = -\cosh\alpha \; ; \tag{7.51}$$

see Hayward (1993). To avoid fixing N_0, one introduces the additional parametrization (Brotz and Kiefer 1997)

$$N_0(t) = \dot{\alpha}(t). \tag{7.52}$$

As with (7.42) and (7.43) above, one must make the replacement

$$(2G)^{-1} \int \mathrm{d}t \; R_0^2 \dot{\alpha} \quad \rightarrow \quad \int \mathrm{d}t \; (\pi_\alpha \dot{\alpha} - N_0 \mathcal{C}_\alpha) \tag{7.53}$$

in the action, obtaining the new constraint

$$\mathcal{C}_\alpha = \pi_\alpha - \frac{A}{8\pi G} \approx 0 \;, \tag{7.54}$$

where $A = 4\pi R_0^2$ is the surface area of the bifurcation sphere. One finds that α and A are canonically conjugate variables; see Carlip and Teitelboim (1995).

The quantum constraints can then be solved and a plane-wave-like solution obtained:

$$\Psi(\alpha, \tau, \lambda) = \chi(M, q) \exp\left[\frac{\mathrm{i}}{\hbar}\left(\frac{A(M,q)\alpha}{8\pi G} - M\tau - q\lambda\right)\right], \tag{7.55}$$

where $\chi(M, q)$ is an arbitrary function of M and q; one can construct superpositions of the solutions (7.55) in the standard way by integrating over M and q.

By varying the phase in (7.55) with respect to M and q, one obtains the classical equations

$$\alpha = 8\pi G \left(\frac{\partial A}{\partial M}\right)^{-1} \tau = \kappa\tau, \tag{7.56}$$

$$\lambda = \frac{\kappa}{8\pi G} \frac{\partial A}{\partial q} \tau = \Phi\tau. \tag{7.57}$$

The solution (7.55) holds for non-extremal holes. If one were to perform a similar quantization for extremal holes on their own, the first term in the exponent of (7.55) would be absent.

An interesting analogy with (7.55) is the plane-wave solution for a free non-relativistic particle,

$$\exp(\mathrm{i}kx - \omega(k)t). \tag{7.58}$$

As in (7.55), the number of parameters is one less than the number of arguments, since $\omega(k) = k^2/2m$. A quantization for extremal holes on their own would correspond to choosing a particular value for the momentum at the classical level, say p_0, and demanding that no dynamical variables (x,p) exist for $p=p_0$. However, the usual way to obtain classical correspondence is not from the plane-wave solution (7.58) but from *wave packets* that are constructed by superposing different wave numbers k. This then yields quantum states that are sufficiently concentrated around classical trajectories such as $x=p_0t/m$.

It therefore seems appropriate to proceed similarly for black holes: construct wave packets for non-extremal holes that are concentrated around the classical values (7.56) and (7.57) and then *extend* them by hand to the extremal limit. This would correspond to 'extremization after quantization', in contrast to the 'quantization after extremization' done above. Expressing M in (7.55) as a function of A and q and using Gaussian weight functions, one has

$$\Psi(\alpha,\tau,\lambda)=\int\limits_{A>4\pi q^2}\mathrm{d}A\,\mathrm{d}q\,\exp\left[-\frac{(A-A_0)^2}{2(\Delta A)^2}-\frac{(q-q_0)^2}{2(\Delta q)^2}\right]$$
$$\times\exp\left[\frac{\mathrm{i}}{\hbar}\left(\frac{A\alpha}{8\pi G}-M(A,q)\tau-q\lambda\right)\right]. \tag{7.59}$$

The result of this calculation is given and discussed in Kiefer and Louko (1999). As expected, one finds Gaussian packets that are concentrated around the classical values (7.56) and (7.57). As in the case of the free particle, the wave packets exhibit dispersion with respect to the Killing time τ. Using for ΔA the Planck length squared, that is, $\Delta A\propto G\hbar\approx 2.6\times 10^{-66}\mathrm{cm}^2$, one finds for the typical dispersion time in the Schwarzschild case

$$\tau_*=\frac{128\pi^2R_{\mathrm{S}}^3}{G\hbar}\approx 10^{65}\left(\frac{M}{M_\odot}\right)^3\text{years}\,. \tag{7.60}$$

Note that this agrees well with the black-hole evaporation time (7.18). The dispersion of the wave packet gives the timescale after which the semiclassical approximation breaks down.

Coming back to the charged case, and going to the extremal limit $\sqrt{G}M=|q|$, one finds that the widths of the wave packet (7.59) are *independent* of τ for large τ. This is due to the fact that $\kappa=0$ for the extremal black hole and therefore no evaporation takes place. If one takes, for example, $\Delta A\propto G\hbar$ and $\Delta q\propto\sqrt{G\hbar}$, one finds for the α-dependence of (7.59) as $\tau\to\infty$ the factor

$$\exp\left(-\frac{\alpha^2}{128\pi^2}\right), \tag{7.61}$$

which is independent of both τ and $\hbar$. It is clear that this packet, although concentrated at the value $\alpha=0$ for extremal holes, has support also for $\alpha\neq 0$ and is qualitatively not different from a wave packet that is concentrated at a value $\alpha\neq 0$ close to extremality.

An interesting question is the possible occurrence of a naked singularity, for which $\sqrt{G}M<|q|$. Certainly, the above boundary conditions do not include the case of a singular three-geometry. However, the wave packets discussed above also contain parameter values that would correspond to the 'naked' case. Such geometries could be avoided if one imposed the boundary condition that the wave function vanishes for such values. But then continuity would force the wave function to vanish also on the boundary, that is, at $\sqrt{G}M = |q|$. This would mean that extremal black holes could not exist at all in quantum gravity—an interesting speculation.

A thermodynamical interpretation of (7.55) can only be obtained if an appropriate transition into the Euclidean regime is performed. This transition is achieved by the 'Wick rotations' $\tau \to -\mathrm{i}\beta\hbar$, $\alpha \to -\mathrm{i}\alpha_{\mathrm{E}}$ (from (7.52), it is clear that α is connected to the lapse function and must be treated like τ), and $\lambda \to -\mathrm{i}\beta\hbar\Phi$. By demanding regularity of the Euclidean line element, one arrives at the conclusion that $\alpha_{\mathrm{E}} = 2\pi$. But this means that the Euclidean version of (7.56) reads $2\pi = \kappa\beta\hbar$, which, with $\beta = (k_{\mathrm{B}}T_{\mathrm{BH}})^{-1}$, becomes exactly the expression for the Hawking temperature (1.34). Alternatively, one could use (1.34) to derive $\alpha_{\mathrm{E}} = 2\pi$.

The Euclidean version of the state (7.55) then reads

$$\Psi_{\mathrm{E}}(\alpha, \tau, \lambda) = \chi(M, q) \exp\left(\frac{A}{4G\hbar} - \beta M - \beta\Phi q\right) . \tag{7.62}$$

We see in the exponent of (7.62) the appearance of the Bekenstein–Hawking entropy. Of course, (7.62) is still a pure state and should not be confused with a partition function. But the factor $\exp[A/(4G\hbar)]$ in (7.62) gives directly the enhancement factor for the rate of black-hole pair creation relative to ordinary pair creation (Hawking and Penrose 1996). It must be emphasized that S_{BH} arises entirely from a boundary term at the horizon ($r \to 0$).

It is now clear that a quantization scheme that treats extremal black holes as a limiting case gives $S_{\mathrm{BH}} = A/(4G\hbar)$ also for the extremal case.[6] This coincides with the result found from string theory; see Section 9.2.5. On the other hand, quantizing extremal holes on their own would yield $S_{\mathrm{BH}} = 0$. From this point of view, it is also clear why the extremal (Kerr) black hole that occurs in the transition from the disc-of-dust solution to the Kerr solution has entropy $A/(4G\hbar)$; see Neugebauer (1998). If $S_{\mathrm{BH}} \neq 0$ for the extremal hole (which has a temperature of zero), the stronger version of the Third Law of Thermodynamics (which would require $S \to 0$ for $T \to 0$) does not hold. This is not particularly disturbing, since many systems in ordinary thermodynamics (such as glasses) violate the strong form of the Third Law; it just means that the system does not approach a unique state for $T \to 0$.

The above discussion has been performed for a pure black hole without inclusion of matter degrees of freedom. In the presence of other variables, it is no longer possible to find simple solutions like (7.55). One possible treatment is to perform a semiclassical approximation as presented in Section 5.4. In this way, one can recover a functional Schrödinger equation for matter fields on a black-hole background. For simple situations, this equation can be solved (Demers and Kiefer 1996). Although the resulting solution is, of course, a pure state, the expectation value of the particle number operator exhibits a Planckian distribution with respect to the Hawking temperature—this is how Hawking radiation is recovered in this approach. For this reason, one might even speculate that the information-loss problem for black holes is not a real problem, since only pure states appear for the full system; see Section 7.6. In fact, the mixed nature

[6]We set $k_{\mathrm{B}} = 1$ here and in the following.

of Hawking radiation can be understood by considering the process of decoherence (Kiefer 2001*a*, 2004*a*). It is even possible that the Bekenstein–Hawking entropy (7.23) could be calculated from the decohering influence of additional degrees of freedom such as the quasi-normal modes of the black hole; cf. the remarks in the next section.

We mention, finally, that loop quantum gravity has also been applied to spherically symmetric systems at the kinematical level (Bojowald 2004). One advantage compared with the full theory is that the flux variables commute with each other, so that there exists a flux representation. It turns out that loop quantization of the reduced model and reduction of the states of the full theory to spherical symmetry lead to the same result. It seems that singularities are avoided (Modesto 2004). This also follows if one uses geometrodynamical variables in connection with methods from loop quantum gravity (Husain and Winkler 2005).

7.3 Black-hole spectroscopy and entropy

7.3.1 Possible spectra for the area

The results of the last subsection indicate that black holes are truly quantum objects. In fact, as Bekenstein (1999) has emphasized, they might play the same role in the development for quantum gravity that atoms played in quantum mechanics. In the light of this possible analogy, one may wonder whether black holes possess a discrete spectrum of states like atoms. Arguments in favour of this idea were given in Bekenstein (1974, 1999); cf. also Mukhanov (1986). Bekenstein (1974) noticed that the horizon area of a (non-extremal) black hole can be treated in the classical theory as a mechanical adiabatic invariant. This is borne out in gedanken experiments in which one shoots charged particles (in the Reissner–Nordström case) or scalar waves (in the Kerr case) with appropriate energies into the hole. From experience with quantum mechanics, one would expect that the corresponding quantum entity would possess a discrete spectrum. The simplest possibility, certainly, is to have a constant spacing between the eigenvalues, that is, $A_n \propto n$ for $n \in \mathbb{N}$. (In the Schwarzschild case, this would entail $M_n \propto \sqrt{n}$ for the mass values.) This can be tentatively concluded, for example, from the 'Euclidean' wave function (7.62): if one imposed an ad hoc Bohr–Sommerfeld quantization rule, one would find, recalling (7.54),

$$2\pi n\hbar = \oint \pi_{\alpha_{\mathrm{E}}}\,\mathrm{d}\alpha_{\mathrm{E}} = \int_0^{2\pi} \frac{A\,\mathrm{d}\alpha_{\mathrm{E}}}{8\pi G} = \frac{A}{4G}. \tag{7.63}$$

(Recall from above that α_{E} ranges from 0 to 2π.) A similar result follows if the range of the time parameter τ in the Lorentzian version is assumed to be compact, as in momentum quantization on finite spaces (Kastrup 1996).

A different argument to fix the factor in the area spectrum goes as follows (Mukhanov 1986; Bekenstein and Mukhanov 1995). One assumes the quantization condition

$$A_n = \alpha l_{\mathrm{P}}^2 n,\ n \in \mathbb{N}, \tag{7.64}$$

with some undetermined constant α. The energy level n will be degenerate with multiplicity $g(n)$, so one would expect the identification

$$S = \frac{A}{4l_{\mathrm{P}}^2} + \text{ constant } = \ln g(n). \tag{7.65}$$

Demanding $g(1) = 1$ (i.e. assuming that the entropy of the ground state vanishes), this leads with (7.64) to

$$g(n) = \mathrm{e}^{\alpha(n-1)/4}.$$

Since this must be an integer, one has the options

$$\alpha = 4\ln k\ , \quad k = 2, 3, \dots\ , \tag{7.66}$$

and thus $g(n) = k^{n-1}$. Note that the spectrum would then differ slightly from (7.63). For information-theoretic reasons ('it from bit'; cf. e.g. Wheeler (1990)) one would prefer the value $k = 2$, leading to $A_n = (4\ln 2)l_{\mathrm{P}}^2 n$.

In the Schwarzschild case, the energy spacing between consecutive levels is found from

$$\Delta A = 32\pi G^2 M \Delta M = (4\ln k)l_{\mathrm{P}}^2$$

to be

$$\Delta M = \Delta E =: \hbar\tilde{\omega}_k = \frac{\hbar \ln k}{8\pi G M}, \tag{7.67}$$

with the fundamental frequency

$$\tilde{\omega}_k = \frac{\ln k}{8\pi G M} = (\ln k)\frac{T_{\mathrm{BH}}}{\hbar}. \tag{7.68}$$

The black-hole emission spectrum would then be concentrated at multiples of this fundamental frequency—unlike the continuous thermal spectrum of Hawking radiation. In fact, one would have a deviation from the Hawking spectrum even for large black holes, that is, black holes with masses $M \gg m_{\mathrm{P}}$. Another consequence would be that quanta with $\omega < \tilde{\omega}_k$ could not be absorbed by the black hole. Note that $\tilde{\omega}_k$ is of the same order as the frequency corresponding to the maximum of the Planck spectrum.

7.3.2 Loop quantum gravity and quasi-normal modes

A discrete spectrum for the black-hole horizon is found in the context of loop quantum gravity (Chapter 6). This is a direct consequence of the area quantization discussed in Section 6.2. However, this spectrum is not equidistant, and thus need not be in conflict with Hawking radiation. Consider the intersection of a spin network with the surface of a black hole. In the case of 'punctures' only (cf. (6.34)), the state is characterized by a set of spins $\{j_i\}$. The dimension of the corresponding boundary (horizon) Hilbert space is thus

$$\prod_{i=1}^{n}(2j_i + 1).$$

The spectrum in this case is given by (6.34). Denoting the minimal spin by $j_{\min}$, the corresponding area value is

$$A_0 := 8\pi\beta l_{\mathrm{P}}^2 \sqrt{j_{\min}(j_{\min}+1)}. \tag{7.69}$$

It has been suggested that the dominating contribution to the entropy comes from the $j_{\rm min}$ contributions (but see below). Following Dreyer (2003), we consider the number of links with spin $j_{\rm min}$,

$$N = \frac{A}{A_0} = \frac{A}{8\pi\beta l_{\rm P}^2\sqrt{j_{\rm min}(j_{\rm min}+1)}}.$$

The number of microstates is then

$$N_{\rm ms} = (2j_{\rm min}+1)^N.$$

This is only equal to the desired result from the Bekenstein–Hawking entropy, that is, equal to $\exp(A/4l_{\rm P}^2)$, if the Barbero–Immirzi parameter has the special value

$$\beta = \frac{\ln(2j_{\rm min}+1)}{2\pi\sqrt{j_{\rm min}(j_{\rm min}+1)}}. \tag{7.70}$$

What is the value of $j_{\rm min}$? If the underlying group is SU(2), as is usually assumed to be the case, one has $j_{\rm min} = 1/2$ and thus $\beta = \ln 2/\pi\sqrt{3}$. This result has also been found in a calculation by Ashtekar *et al.* (1998) where it was assumed that the degrees of freedom are given by a Chern–Simons theory on the horizon.

An important concept in this context is the 'isolated horizon' of a black hole; cf. Ashtekar *et al.* (2000). An isolated horizon is a generalization of an event horizon to non-stationary black holes. It is a local concept and does not need to admit a Killing field. It is needed here for the derivation of the entropy in order to distinguish the surface of a black-hole horizon from any other surface for which the spectrum (6.34) holds.

Taking SO(3) as the underlying group instead, one has $j_{\rm min} = 1$ and thus $\beta = \ln 3/\pi\sqrt{2}$. If a link with spin $j_{\rm min}$ is absorbed or created at the black-hole horizon, the area changes by

$$\Delta A = A_0 = 8\pi\beta l_{\rm P}^2\sqrt{j_{\rm min}(j_{\rm min}+1)}, \tag{7.71}$$

which is equal to $4(\ln 2)l_{\rm P}^2$ in the SU(2) case and $4(\ln 3)l_{\rm P}^2$ in the SO(3) case. In the SU(2) case, the result for ΔA corresponds to that advocated by Bekenstein and Mukhanov (1995), although the spectrum here is not equidistant (which is why there is no conflict here with the spectrum of the Hawking radiation for large mass).

The situation is, however, not so simple. As was demonstrated by Domagala and Lewandowski (2004), spins greater than the minimal spin are not negligible and therefore have to be taken into account when the entropy is calculated. Therefore the entropy obtained in the earlier papers, which led to the result (7.70) for β, was an underestimate. If one demands again that in the highest order the result is equal to the Bekenstein–Hawking entropy, the equation which fixes the Barbero–Immirzi parameter is then

$$2\sum_{n=1}^{\infty}\exp\left(-2\pi\beta\sqrt{\frac{n(n+2)}{4}}\right) = 1,$$

which can only be solved numerically, with the result (Meissner 2004)

$$\beta = 0.23753295796592\ldots. \tag{7.72}$$

A detailed description of the computation of entropy in loop quantum gravity has been presented in Agullo *et al.* (2010).

If one nevertheless took the value $k = 3$ in (7.67), one would find a change in the mass given by

$$\Delta M = \frac{\ln 3\; m_{\mathrm{P}}^2}{8\pi M}, \tag{7.73}$$

corresponding to the fundamental frequency

$$\tilde{\omega}_3 = \frac{\ln 3}{8\pi GM} \equiv \frac{\ln 3\; T_{\mathrm{BH}}}{\hbar} \approx 8.85 \frac{M_{\odot}}{M}\ \mathrm{kHz}. \tag{7.74}$$

It was emphasized by Dreyer (2003) that $\tilde{\omega}_3$ coincides with the real part of the asymptotic frequency for the quasi-normal modes of the black hole. These modes are characteristic oscillations of the black hole before it settles into its stationary state; see, for example, Kokkotas and Schmidt (1999) for a review. As was conjectured by Hod (1998) on the basis of numerical evidence and shown by Motl (2003), the frequency of the quasi-normal modes in the limit $n \to \infty$ is

$$\begin{aligned} \omega_n &= -\frac{\mathrm{i}(n+\frac{1}{2})}{4GM} + \frac{\ln 3}{8\pi GM} + \mathcal{O}\left(n^{-1/2}\right) \\ &= -\mathrm{i}\kappa\left(n + \frac{1}{2}\right) + \frac{\kappa}{2\pi}\ln 3 + \mathcal{O}\left(n^{-1/2}\right)\ ; \end{aligned} \tag{7.75}$$

see also Neitzke (2003). The imaginary part indicates that one is dealing with damped oscillations. It might be that the black-hole entropy arises from the quantum entanglement between the black hole and the quasi-normal modes (Kiefer 2004*a*). The quasi-normal modes would then serve as an environment leading to decoherence (Section 10.2). This, however, would still have to be shown. Interestingly, at least for the Schwarzschild black hole, a quantum measurement of the quasi-normal modes would introduce a minimal noise temperature that is exactly equal to the Hawking temperature (Kiefer 2004*b*). If one had performed this analysis before the advent of Hawking's work, one would have concluded that there is a temperature associated with the real part in (7.75), which is proportional to $\hbar$ and equal to (1.35).

The imaginary part of the frequency (7.75) in this limit is equidistant in n. This could indicate an intimate relation with Euclidean quantum gravity and explain why the Euclidean version readily provides expressions for the black-hole temperature and entropy: if, in the Euclidean theory, one considered a wave function of the form

$$\psi_{\mathrm{E}} \sim \mathrm{e}^{\mathrm{i}n\kappa t_{\mathrm{E}}},$$

one would have to demand that the Euclidean time t_{E} be periodic with period $8\pi GM$. This is, however, precisely the inverse of the Hawking temperature, in accordance with the result that the Euclidean time must have this periodicity if the line element is to be regular (see e.g. Hawking and Penrose 1996).

7.3.3 Logarithmic corrections to the entropy

Proceeding in loop quantum gravity to the next order in the calculation, one arrives at the following value for the entropy:

$$S = \frac{A}{4l_{\rm P}^2} - \frac{1}{2}\ln\left(\frac{A}{l_{\rm P}^2}\right) + \mathcal{O}(1). \tag{7.76}$$

The logarithmic correction term is, in fact, independent of β. Correction terms of this form (mostly with coefficients $-1/2$ or $-3/2$) have also been found in other approaches; see Page (2005) for a review.

The occurrence of such logarithmic terms can be understood using simple models (Kiefer and Kolland 2008). Let us consider a set of N spin-1/2 particles, of which n point up and $N-n$ point down. Since N is assumed to be given, we have a microcanonical ensemble. We define the entropy as the logarithm of the number of configurations with n spins up and $N-n$ spins down,

$$S = \ln\binom{N}{N-n} = \ln\binom{N}{n}. \tag{7.77}$$

In a concrete example, n might correspond to the magnetization as the fixed macroscopic quantity.

Let us consider first the 'equilibrium case' $n = N/2$. Using Stirling's formula

$$\ln N! = N\ln N - N + \frac{1}{2}\ln N + \frac{1}{2}\ln(2\pi) + \mathcal{O}\left(\frac{1}{N}\right),$$

one gets from (7.77), after terms of order $1/N$ have been neglected,

$$S = N\ln 2 - \frac{1}{2}\ln N + \frac{1}{2}\ln\frac{2}{\pi}. \tag{7.78}$$

Defining $S_0 := N\ln 2$, we see that the relative contribution of the second term is just of order $\ln N/N$. We can then write, approximately,

$$S \approx S_0 - \frac{1}{2}\ln S_0,$$

and thus find a logarithmic correction term as in (7.76).

In the general case we get (assuming both n and N to be large numbers)

$$S = -N(w\ln w + (1-w)\ln(1-w)) - \frac{1}{2}\ln(Nw(1-w)) - \frac{1}{2}\ln(2\pi), \tag{7.79}$$

where $w = n/N$. Defining now

$$S_0 := -N\left(w\ln w + (1-w)\ln(1-w)\right), \tag{7.80}$$

we find

$$S = S_0 - \frac{1}{2}\ln S_0 - \frac{1}{2}\ln\left(\frac{2\pi w(1-w)}{\alpha}\right), \tag{7.81}$$

where $\alpha = (w-1)\ln(1-w) - w\ln w$.

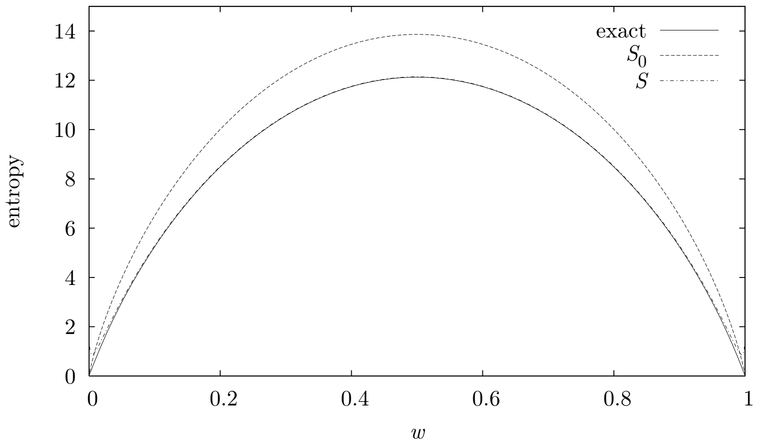

Fig. 7.4 Comparison of the exact entropy (7.77) with the approximation including the logarithmic correction term, (7.81), and the approximation without it, (7.80), for $N = 20$. The difference between the exact expression and (7.81) is hardly noticeable. Reprinted from Kiefer and Kolland (2008) with kind permission from Springer Science+Business Media B.V.

Figure 7.4 shows a comparison of the exact expression (7.77) for the entropy with the approximations (7.80) and (7.81). One can easily see that (7.81) is an excellent approximation unless n or N becomes small.

The origin of the logarithmic corrections to the Bekenstein–Hawking entropy can thus be understood from statistical mechanics; for example, by considering the application of Stirling's formula to higher orders. These corrections are thus not necessarily of quantum origin, as can be seen from the non-occurrence of $\hbar$ in the above expressions. A subtle issue is whether the counting should treat the microscopic entities ('particles') as distinguishable or as indistinguishable (Kiefer and Kolland 2008). The first case would correspond to the application of classical Boltzmann statistics, while the second case would correspond to Bose statistics. Loop quantum gravity employs the former, while string theory uses the latter. Why this is so is far from clear.

7.4 Quantum theory of collapsing dust shells

In this section, a particular model will be described in some detail, but without too many technicalities. This concerns the collapse of a null dust shell. In the classical theory, the collapse leads to the formation of a black hole. We shall see that it is possible to construct an exact quantum theory of this model in which the dynamical evolution is *unitary* with respect to asymptotic observers (since one has an asymptotically flat space, a semiclassical time exists, which is just the Killing time at asymptotic infinity). As a consequence of the unitary evolution, the classical singularity is fully avoided in the quantum theory: if the collapsing shell is described by a wave packet, the evolution leads to a superposition of a black-hole and a white-hole horizon, yielding a wave function that vanishes at zero radial coordinate. At late times, only an expanding wave packet (an expanding shell) is present. A detailed exposition of this model can be found in Hájíček (2003).

7.4.1 Covariant gauge fixing

A problem of principle that arises is the need to represent the expanding and collapsing shells on the same background manifold. What does this mean? Following Section 2 of Hájíček and Kiefer (2001*a*), let $(\mathcal{M}, g)$ be a globally hyperbolic space–time,

$$\mathcal{M} = \Sigma \times \mathbb{R}.$$

The manifold $\mathcal{M}$ is called the 'background manifold' and is uniquely determined for a given three-manifold Σ (the 'initial-data manifold'). The four-dimensional diffeomorphism group $\mathrm{Diff}\,\mathcal{M}$ is often considered as the 'gauge group' of GR; it pushes the points of $\mathcal{M}$ around, so points are not a gauge-invariant concept. In fact, usually only a subgroup of $\mathrm{Diff}\,\mathcal{M}$ plays the role of a gauge group (describing 'redundancies', in the language of Section 4.2.5). Everything else describes physically relevant symmetries, for example asymptotic rotations in an asymptotically flat space.

The group $\mathrm{Diff}\,\mathcal{M}$ acts on $\mathrm{Riem}\,\mathcal{M}$, the space of all Lorentzian metrics on $\mathcal{M}$. A particular representative metric for each geometry on $\mathcal{M}$ (in some open set) is chosen by 'covariant gauge fixing', that is, by the choice of a section σ,

$$\sigma : \mathrm{Riem}\,\mathcal{M}/\mathrm{Diff}\,\mathcal{M} \mapsto \mathrm{Riem}\,\mathcal{M}.$$

Thereby, points are defined by coordinates on a fixed manifold. As has already been emphasized in Section 4.2.3, a transformation between two covariant gauge fixings σ and σ' is not a single diffeomorphism, but forms a much larger group (Bergmann and Komar 1972); it corresponds to one coordinate transformation for each solution of the field equations, which is different from a single coordinate transformation on the background manifold. Hájíček and Kijowski (2000) have shown that, given a section σ, one can construct a map from $\mathrm{Riem}\,\mathcal{M}/\mathrm{Diff}\,\mathcal{M} \times \mathrm{Emb}(\Sigma, \mathcal{M})$, where $\mathrm{Emb}(\Sigma, \mathcal{M})$ is the space of embeddings of the initial data surface Σ into $\mathcal{M}$, to the ADM phase space Γ of GR. (This works only if the evolved space–times do not admit an isometry.) It was shown that this map is invertible and can be extended to neighbourhoods of Γ and $\mathrm{Riem}\,\mathcal{M}/\mathrm{Diff}\,\mathcal{M} \times \mathrm{Emb}(\Sigma, \mathcal{M})$. In this way, a transformation from ADM variables to embedding variables (see Section 5.2.1) is implemented. The use of embedding variables was called 'Kuchař decomposition' by Hájíček and Kijowski (2000).

The Schwarzschild case may serve as a simple illustration. The transformation between Kruskal and Eddington–Finkelstein coordinates is *not* a coordinate transformation on a background manifold, because this transformation is solution dependent (it depends on the mass M of the chosen Schwarzschild solution). It thus represents a *set* of coordinate transformations, one for each M, and is thus a transformation between different gauge fixings σ and σ'. One background manifold is obtained if one identifies all points with the same values of the Kruskal coordinates, and a different background manifold results if one identifies all points with the same Eddington–Finkelstein coordinates.

7.4.2 Embedding variables for the classical theory

Here we consider the dynamics of a (spherically symmetric) null dust shell in GR. In this subsection, we shall first identify appropriate coordinates on a background manifold and then perform the explicit transformation to embedding variables. This will then serve as the natural starting point for the quantization in the next subsection.

Any classical solution describing this system has a simple structure: inside the shell, the space–time is flat, whereas outside, it is isometric to a part of the Schwarzschild metric with mass M. The two geometries must match along a spherically symmetric null hypersurface describing the shell. All physically distinct solutions can be labelled by three parameters: $\eta \in \{-1, +1\}$, distinguishing between the outgoing ($\eta = +1$) and ingoing ($\eta = -1$) null surfaces; the asymptotic time of the surface, that is, the retarded time $u = T - R \in (-\infty, \infty)$ for $\eta = +1$, and the advanced time $v = T + R \in (-\infty, \infty)$ for $\eta = -1$; and the mass $M \in (0, \infty)$. An ingoing shell creates a black-hole (event) horizon at $R = 2M$ and ends up in the singularity at $R = 0$. The outgoing shell starts from the singularity at $R = 0$ and emerges from a white-hole (particle) horizon at $R = 2M$. We shall follow Hájíček and Kiefer (2001*a,b*); see also Hájíček (2003).

The Eddington–Finkelstein coordinates do not define a covariant gauge fixing, since it turns out that there are identifications of points in different solutions that do not keep an asymptotic family of observers fixed. Instead, double-null coordinates U and V are chosen on the background manifold $\mathcal{M} = \mathbb{R}_+ \times \mathbb{R}$ (which is effectively two-dimensional due to spherical symmetry). In these coordinates (which will play the role of the embedding variables), the metric has the form

$$\mathrm{d}s^2 = -A(U,V)\,\mathrm{d}U\,\mathrm{d}V + R^2(U,V)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2). \tag{7.82}$$

From the demand that the metric be regular at the centre and continuous at the shell, the coefficients A and R are uniquely defined for any physical situation defined by the variables M (the energy of the shell), η, and w (the location of the shell, where $w = u$ for the outgoing and $w = v$ for the ingoing case).

Let us consider first the case $\eta = 1$. In the Minkowski part of the solution, $U > u$, one finds

$$A = 1\,, \quad R = \frac{V-U}{2}. \tag{7.83}$$

In the Schwarzschild part, $U < u$,

$$R = 2M\kappa\left(\left(\frac{V-u}{4M} - 1\right)\exp\left(\frac{V-U}{4M}\right)\right) =: 2M\kappa(f_+), \tag{7.84}$$

where κ is the 'Kruskal function' (not to be confused with the surface gravity κ), which is defined by its inverse as

$$\kappa^{-1}(y) = (y-1)e^y, \tag{7.85}$$

and

$$A = \frac{1}{\kappa(f_+)e^{\kappa(f_+)}}\frac{V-u}{4M}\exp\left(\frac{V-U}{4M}\right). \tag{7.86}$$

With these expressions, one can verify that A and R are continuous at the shell, as required. We note that these expressions contain u as well as M, and these will become conjugate variables in the canonical formalism.

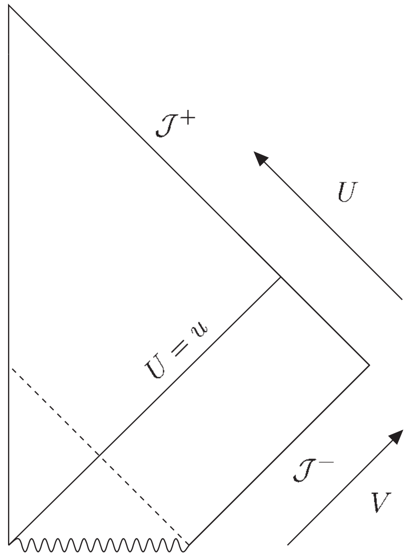

Fig. 7.5 Penrose diagram for the outgoing shell in the classical theory. The shell is at $U = u$.

The Penrose diagram for the outgoing shell is shown in Fig. 7.5. It is important to note that the background manifold possesses a unique asymptotic region, with $\mathcal{J}^-$ defined by $U \to -\infty$ and $\mathcal{J}^+$ by $V \to +\infty$.

In the case of an ingoing shell ($\eta = -1$), one finds (7.83) again for $V < v$, and for $V > v$ one finds

$$R = 2M\kappa(f_-), \quad A = \frac{1}{\kappa(f_-)e^{\kappa(f_-)}} \frac{v-U}{4M} \exp\left(\frac{V-U}{4M}\right), \tag{7.87}$$

where

$$f_- := \left(\frac{v-U}{4M} - 1\right) \exp\left(\frac{V-U}{4M}\right).$$

These expressions result from (7.86) by the substitution $V - u \to v - U$.

As a result of the gauge fixing, the set of solutions (η, M, w) can be written as a set of (η, M, w)-dependent metric fields (7.82) and a set of shell trajectories on a *fixed* background manifold $\mathcal{M}$. Here, the corresponding functions A and R have the form

$$A(\eta, M, w; U, V)\ , \quad R(\eta, M, w; U, V), \tag{7.88}$$

and the trajectory of the shell on the background manifold is simply $U = u$ for $\eta = +1$ and $V = v$ for $\eta = -1$.

The next step is the explicit transformation to embedding variables (the Kuchař decomposition). The standard (ADM) formulation of the shell was studied in Louko *et al.* (1998); see also Kraus and Wilczek (1995). The spherically symmetric metric is written in the form

$$\mathrm{d}s^2 = -N^2\,\mathrm{d}\tau^2 + L^2(\mathrm{d}\rho + N^\rho\,\mathrm{d}\tau)^2 + R^2\,\mathrm{d}\Omega^2, \tag{7.89}$$

and the shell is described by its radial coordinate $\rho = \mathbf{r}$. The action reads

$$S_0 = \int \mathrm{d}\tau \left[\mathbf{p}\dot{\mathbf{r}} + \int \mathrm{d}\rho \, (P_L \dot{L} + P_R \dot{R} - H_0) \right], \tag{7.90}$$

and the Hamiltonian is

$$H_0 = N\mathcal{H}_\perp + N^\rho \mathcal{H}_\rho \ + N_\infty E_\infty,$$

where $N_\infty := \lim_{\rho\to\infty} N(\rho)$, E_∞ is the ADM mass, and N and N^ρ are the lapse and shift functions. The constraints read

$$\mathcal{H}_\perp = \frac{LP_L^2}{2R^2} - \frac{P_L P_R}{R} + \frac{RR''}{L} - \frac{RR'L'}{L^2} + \frac{R'^2}{2L} - \frac{L}{2} + \frac{\eta \mathbf{p}}{L}\delta(\rho - \mathbf{r}) \approx 0, \tag{7.91}$$

$$\mathcal{H}_\rho = P_R R' - P_L' L - \mathbf{p}\delta(\rho - \mathbf{r}) \approx 0, \tag{7.92}$$

where the primes and dots denote derivatives with respect to ρ and τ, respectively. These are the same constraints as in (7.28) and (7.30), except for the contribution of the shell.

The task is to transform the variables in the action S_0. This transformation will be split into two steps. The first step is a transformation of the canonical coordinates $\mathbf{r}$, $\mathbf{p}$, L, P_L, R, and P_R *on the constraint surface* Γ defined by the constraints (7.91) and (7.92). The new coordinates are u and $p_u = -M$ for $\eta = +1$, and v and $p_v = -M$ for $\eta = -1$, and the embedding variables $U(\rho)$ and $V(\rho)$.

The second step is an *extension* of the functions u, v, p_u, p_v, $U(\rho)$, $P_U(\rho)$, $V(\rho)$, and $P_V(\rho)$ *off the constraint surface*, where the functions u, v, p_u, p_v, $U(\rho)$, and $V(\rho)$ are defined by the above transformation, and $P_U(\rho)$, $P_V(\rho)$ by $P_U(\rho)|_\Gamma = P_V(\rho)|_\Gamma = 0$. The extension must satisfy the condition that the functions form a canonical chart in a neighbourhood of Γ. That such an extension exists was shown by Hájíček and Kijowski (2000). The details of the calculation can be found in Hájíček and Kiefer (2001*a*). The result is the action

$$S = \int d\tau \, (p_u \dot{u} + p_v \dot{v} - n p_u p_v) + \int d\tau \int_0^\infty d\rho \, (P_U \dot{U} + P_V \dot{V} - H), \tag{7.93}$$

where $H = N^U P_U + N^V P_V$, and n, $N^U(\rho)$, and $N^V(\rho)$ are Lagrange multipliers. The first term in (7.93) contains the physical variables (observables), while the second term contains the gauge variables. Both classical solutions are contained in the single constraint

$$p_u p_v = 0, \tag{7.94}$$

that is, one has $p_v = 0$ for $\eta = 1$ and $p_u = 0$ for $\eta = -1$. Observe that the Poisson algebra of the chosen set of observables p_u and u for $\eta = +1$ and p_v and v for $\eta = -1$ is gauge-invariant in spite of the fact that it has been obtained by a calculation based on a gauge choice (the double-null coordinates U and V). Therefore, the quantum theory will also be gauge-invariant. A crucial point is that the new phase space has non-trivial boundaries,

$$p_u \leq 0 \;, \quad p_v \leq 0, \quad \frac{-u+v}{2} > 0. \tag{7.95}$$

The boundary defined by the last inequality is due to the classical singularity. The system has now been brought into a form that can be taken as the starting point for quantization.

7.4.3 Quantization

The task is to quantize the physical degrees of freedom defined by the action

$$S_{\text{phys}} = \int \mathrm{d}\tau \, (p_u \dot{u} + p_v \dot{v} - n p_u p_v) \;; \tag{7.96}$$

cf. (7.93). The appropriate method is *group quantization*; see e.g. Isham (1984). This method is suited in particular to implementing conditions such as (7.95). It is based on the choice of a set of Dirac observables forming a Lie algebra. This algebra generates a group of transformations which respects all boundaries and thus ensures that information about such boundaries is implemented in the quantum theory. The method automatically leads to self-adjoint operators for the observables. One obtains in particular a self-adjoint Hamiltonian and, consequently, a unitary dynamics.

The application of this method to the null-dust shell was presented in detail in Hájíček (2001, 2003). A complete system of Dirac observables is given by p_u, p_v, $D_u := up_u$, and $D_v := vp_v$. Thus, they commute with the constraint $p_u p_v$. The only non-vanishing Poisson brackets are

$$\{D_u, p_u\} = p_u \;, \quad \{D_v, p_v\} = p_v. \tag{7.97}$$

The Hilbert space is constructed from complex functions $\psi_u(p)$ and $\psi_v(p)$, where $p \in [0, \infty)$. The scalar product is defined by

$$(\psi_u, \phi_u) := \int_0^\infty \frac{\mathrm{d}p}{p} \; \psi_u^*(p)\phi_u(p) \tag{7.98}$$

(and similarly for $\psi_v(p)$). To handle the inequalities (7.95), it is useful to perform the following canonical transformation:

$$t = (u+v)/2, \qquad r = (-u+v)/2, \tag{7.99}$$

$$p_t = p_u + p_v, \qquad p_r = -p_u + p_v. \tag{7.100}$$

The constraint function then assumes the form $p_u p_v = (p_t^2 - p_r^2)/4$. Upon quantization, one obtains the operator $-\hat{p}_t$, which is self-adjoint and has a positive spectrum, $-\hat{p}_t \varphi(p) = p\varphi(p)$, $p \geq 0$. It is the generator of time evolution and corresponds to the energy operator $E := M$. Since r is not a Dirac observable, it cannot be directly transformed into a quantum observable. It turns out that the following construction is useful:

$$\hat{r}^2 := -\sqrt{p}\frac{\mathrm{d}^2}{\mathrm{d}p^2}\frac{1}{\sqrt{p}}. \tag{7.101}$$

This is essentially a Laplacian and corresponds to a concrete choice of factor ordering. It is a symmetric operator that can be extended to a self-adjoint operator. In this process, one is naturally led to the following eigenfunctions of $\hat{r}^2$:

$$\psi(r,p) := \sqrt{\frac{2p}{\pi}}\sin rp\ , \quad r \geq 0. \tag{7.102}$$

One can also construct an operator $\hat{\eta}$ that classically would correspond to the direction of motion of the shell.

The formalism has now reached a stage at which one can start to study concrete physical applications. Of particular interest is the representation of the shell by a narrow wave packet. One takes at $t = 0$ the following family of wave packets:

$$\psi_{\kappa\lambda}(p) := \frac{(2\lambda)^{\kappa+1/2}}{\sqrt{(2\kappa)!}}p^{\kappa+1/2}\mathrm{e}^{-\lambda p}, \tag{7.103}$$

where κ is a positive integer, and λ is a positive number with dimensions of length. By an appropriate choice of these constants, one can prescribe the expectation value of the energy and its variation. A sufficiently narrow wave packet can thus be constructed.

One can show that the wave packets are normalized and that they obey $\psi_{\kappa\lambda}(p) = \psi_{\kappa 1}(\lambda p)$ ('scale invariance'). The expectation value of the energy is calculated as

$$\langle E\rangle_{\kappa\lambda} := \int_0^\infty \frac{\mathrm{d}p}{p}\ p\psi^2_{\kappa\lambda}(p), \tag{7.104}$$

with the result

$$\langle E\rangle_{\kappa\lambda} = \frac{\kappa+1/2}{\lambda}. \tag{7.105}$$

Similarly, one finds for the variance

$$\Delta E_{\kappa\lambda} = \frac{\sqrt{2\kappa+1}}{2\lambda}. \tag{7.106}$$

Since the time evolution of the packet is generated by $-\hat{p}_t$, one has

$$\psi_{\kappa\lambda}(t,p) = \psi_{\kappa\lambda}(p)\mathrm{e}^{-\mathrm{i}pt}. \tag{7.107}$$

More interesting is the evolution of the wave packet in the r-representation. This is given by the integral transform (7.98) of $\psi_{\kappa\lambda}(t,p)$ with respect to the eigenfunctions (7.102). It leads to the exact result

$$\Psi_{\kappa\lambda}(t,r) = \frac{1}{\sqrt{2\pi}}\frac{\kappa!(2\lambda)^{\kappa+1/2}}{\sqrt{(2\kappa)!}}\left[\frac{\mathrm{i}}{(\lambda+\mathrm{i}t+\mathrm{i}r)^{\kappa+1}} - \frac{\mathrm{i}}{(\lambda+\mathrm{i}t-\mathrm{i}r)^{\kappa+1}}\right]. \tag{7.108}$$

One interesting consequence can be immediately deduced:

$$\lim_{r \to 0} \Psi_{\kappa\lambda}(t, r) = 0. \tag{7.109}$$

This means that the probability of finding the shell at vanishing radius is zero! In this sense, the singularity is avoided in the quantum theory. It must be emphasized that this is not a consequence of a certain boundary condition—it is a consequence of the *unitary evolution*. If the wave function vanishes at $r = 0$ for $t \to -\infty$ (the asymptotic condition of an ingoing shell), it will continue to vanish at $r = 0$ for all times. It follows from (7.108) that the quantum shell bounces and re-expands. Hence, no absolute event horizon can form, in contrast to the classical theory. The resulting object might still be indistinguishable from a black hole due to the huge time delay from the gravitational redshift—the re-expansion would be visible from afar only in the far future. Similar features follow from a model by Frolov and Vilkovisky (1981), who consider a null shell in the case where 'loop effects' in the form of Weyl curvature terms are taken into account.

Of interest also is the expectation value of the shell radius; see Hájíček (2001, 2003) for details. Again one recognizes that the quantum shell always bounces and re-expands. An intriguing feature is that an essential part of the wave packet can even be squeezed below the expectation value of its Schwarzschild radius. The latter is found from (7.104) to be (reinserting G)

$$\langle R_0 \rangle_{\kappa\lambda} := 2G\langle E \rangle_{\kappa\lambda} = (2\kappa + 1)\frac{l_{\mathrm{P}}^2}{\lambda}, \tag{7.110}$$

while its variance follows from (7.106),

$$\Delta(R_0)_{\kappa\lambda} = 2G\,\Delta E_{\kappa\lambda} = \sqrt{2\kappa + 1}\frac{l_{\mathrm{P}}^2}{\lambda}. \tag{7.111}$$

The main part of the wave packet is squeezed below the Schwarzschild radius if

$$\langle r \rangle_{\kappa\lambda} + (\Delta r)_{\kappa\lambda} < \langle R_0 \rangle_{\kappa\lambda} - \Delta(R_0)_{\kappa\lambda}.$$

It turns out that this can be achieved either if $\lambda \approx l_{\mathrm{P}}$ (and $\kappa > 2$) or, for bigger λ, if κ is larger by a factor of $(\lambda/l_{\mathrm{P}})^{4/3}$. The wave packet can thus be squeezed below its Schwarzschild radius if its energy is greater than the Planck energy—a genuine quantum effect.

How can this behaviour be understood? The unitary dynamics ensures that the ingoing quantum shell develops into a *superposition* of ingoing and outgoing shells if it reaches the region where a singularity would form in the classical theory. In other words, the singularity is avoided by destructive interference in the quantum theory. This is similar to the quantum-cosmological example of Kiefer and Zeh (1995), in which a superposition of a black hole with a white hole leads to a singularity-free quantum universe; cf. Section 10.2. Here also, the horizon becomes a superposition of a 'black hole' and a 'white hole'—its 'grey' nature can be characterized by the expectation value of the operator $\hat{\eta}$ (a black-hole horizon would correspond to the value -1 and a white-hole horizon to +1). In this scenario, no information-loss paradox would ever arise if such a behaviour occurred for all collapsing matter (which sounds reasonable). In the same way, the principle of cosmic censorship would be implemented, since no naked singularities (in fact, no singularities at all) would form.

7.5 The Lemaître–Tolman–Bondi model

In the last section, we studied the dynamics of a collapsing (null) dust *shell*. The next level of complexity is a self-gravitating inhomogeneous (time-like) dust *cloud*, while keeping spherical symmetry. This is what we shall discuss in this section. The model goes back to Lemaître (1933), who used it as an inhomogeneous model for cosmology, and was later elaborated by Tolman, Bondi, and others; see Plebański and Krasiński (2006) for details and references. For this reason, we shall call it the Lemaître–Tolman–Bondi (LTB) model. It is used today in classical cosmology as a means to understand the formation of voids and structure.

The canonical formulation of the LTB model and and its quantization in the spirit of Section 7.2 were performed in Vaz *et al.* (2001) for a special case (the 'marginal case') and in Kiefer *et al.* (2006) for the generic (including the 'non-marginal') case. The extension to a positive cosmological constant can be found in Franzen *et al.* (2010), and that to a negative cosmological constant in Vaz *et al.* (2008*b*). We shall first introduce the classical model and then turn to its quantization.

7.5.1 The classical LTB model

The LTB model describes a self-gravitating dust cloud. Its energy–momentum tensor is $T_{\mu\nu} = \epsilon(\tau, \rho)u_\mu u_\nu$, where $u^\mu = u^\mu(\tau, \rho)$ is the four-velocity vector of a dust particle with proper time τ and label ρ (ρ thus labels the various shells that together form the dust cloud). The line element for the LTB space–time is given by

$$\mathrm{d}s^2 = -\mathrm{d}\tau^2 + \frac{(\partial_\rho R)^2}{1+2E(\rho)}\mathrm{d}\rho^2 + R^2(\rho)\,\mathrm{d}\Omega^2. \tag{7.112}$$

Inserting this expression into the Einstein equations leads to

$$8\pi G\epsilon(\tau, \rho) = \frac{\partial_\rho F}{R^2\partial_\rho R} \quad \text{and} \quad (\partial_\tau R)^2 = \frac{F}{R} + \frac{\Lambda R^2}{3} + 2E\,, \tag{7.113}$$

where $F(\rho)$ is a non-negative function with the dimensions of a length, and Λ is the cosmological constant. A collapsing dust cloud is described by $\partial_\tau R(\tau, \rho) < 0$. We see from (7.112) that we have to demand that $E > 1/2$ in order for the metric to have the right signature.

There still exists the freedom to rescale the shell index ρ. This can be fixed by demanding

$$R(0, \rho) = \rho\,, \tag{7.114}$$

so that for $\tau = 0$, the label coordinate ρ is equal to the curvature radius R. Now we can express the functions $F(\rho)$ and $E(\rho)$ in terms of the energy density ϵ at $\tau = 0$. From (7.113) one gets

$$F(\rho) = 8\pi G \int_0^{\rho} \epsilon(0, \tilde{\rho})\, \tilde{\rho}^2\, \mathrm{d}\tilde{\rho}, \tag{7.115}$$

$$E(\rho) = \frac{1}{2}[\partial_\tau R(0, \rho)]^2 - \frac{F(\rho)}{2\rho} - \frac{\Lambda\rho^2}{6}\ . \tag{7.116}$$

The interpretation of these quantities is that $F(\rho)/2G \equiv M(\rho)$ is the active gravitating mass inside $R(\tau, \rho)$, while $E(\rho)$ is the total energy of shell ρ. One can also show that $-E$ is a measure of the curvature of the subspaces with constant time. The marginally bound models are those with $E(\rho) = 0$. Here we consider the general case, which includes the non-marginal case $E(\rho) \neq 0$.

The metric (7.112) contains two important special cases. First, for $\partial_\rho F = 0$, we obtain the Schwarzschild metric, $F = 2GM = \text{constant}$.[7] Second, in the limit of homogeneity, one arrives at the Friedmann–Lemaître (FL) equations of standard cosmology; cf. Section 8.1.2. Let us have a brief look at this limit.

Setting $R(\tau, \rho) = \rho a(t)$, where a denotes the scale factor of the FL universe, the second equation in (7.113) reads

$$\dot{a}^2 = \frac{2GM(\rho)}{\rho^3 a} + \frac{\Lambda a^2}{3} + \frac{2E(\rho)}{\rho^2}. \tag{7.117}$$

The corresponding FL equation reads

$$\dot{a}^2 = \frac{8\pi G}{3} a^2 \epsilon + \frac{\Lambda a^2}{3} - k.$$

For dust, we have

$$\epsilon a^3 = \text{constant} =: \frac{3M_0}{4\pi},$$

and thus

$$\dot{a}^2 = \frac{2GM_0}{a} + \frac{\Lambda a^2}{3} - k. \tag{7.118}$$

Comparison of (7.118) with (7.117) gives

$$M(\rho) = M_0 \rho^3 \quad \text{and} \quad E(\rho) = -\frac{k\rho^2}{2}.$$

In fact, $E/M^{2/3} = \text{constant}$ is the invariant definition of the FL limit (Plebański and Krasiński 2006). With these identifications, the LTB metric (7.112) then assumes the standard form of the Robertson–Walker line element,

$$\mathrm{d}s^2 = -\mathrm{d}\tau^2 + a^2(\tau)\left[\frac{\mathrm{d}\rho^2}{1 - k\rho^2} + \rho^2\, \mathrm{d}\Omega^2\right].$$

We see here explicitly the direct correspondence between $-E$ and k.

[7] To see this from (7.115), one has to add an integration constant there or consider a δ-contribution at the origin.

Let us turn from the FL limit back to the LTB model. The equations (7.113) can be solved exactly. While the solutions for $\Lambda = 0$ can be written in terms of elementary functions, the solutions for $\Lambda \neq 0$ entail elliptic functions (Plebański and Krasiński 2006). In the following, we shall restrict ourselves to vanishing Λ and return to the general case at the end.

The solutions show that at the dust proper time $\tau = \tau_0(\rho)$, the shell ρ reaches a curvature radius $R = 0$, that is, the physical singularity. The model thus exhibits a non-simultaneous big-bang (or big-crunch) surface. The parameter τ can only take values between $-\infty$ and $\tau_0(\rho)$. Apart from the freedom in the radial coordinate, the three functions $F(\rho)$, $E(\rho)$, and $\tau_0(\rho)$ determine the model completely.

As in Sections 7.2 and 7.4, we start with the general ansatz (7.26) for a spherically symmetric line element. The gravitational part of the action is

$$S^{\mathrm{g}} = \int \mathrm{d}t \int_0^\infty \mathrm{d}r \left(P_L \dot{L} + P_R \dot{R} - N\mathcal{H}^{\mathrm{g}} - N^r \mathcal{H}_r^{\mathrm{g}} \right) + S_{\partial\Sigma} \,, \tag{7.119}$$

where the Hamiltonian and the diffeomorphism (momentum) constraint are given by (7.28) and (7.30), respectively, and the boundary action $S_{\partial\Sigma}$ is discussed below.

The total action is the sum of (7.119) and an action S^{d} describing the dust. But how do we describe dust in the canonical formalism? Brown and Kuchař (1995) have proposed the following action for dust:

$$S^{\mathrm{d}} = -\tfrac{1}{2} \int \mathrm{d}^4x \, \sqrt{-g}\epsilon(x) \left[g_{\alpha\beta} u^\alpha u^\beta + 1 \right] ,$$

where the u^α are the components of the four-velocity. Writing

$$u_\alpha = -\tau_{,\alpha} + W_k Z^k_{,\alpha},$$

the dust is described by eight space–time scalars: ϵ (the proper energy density of the dust), τ (the proper time along the flow lines), Z^k (the comoving coordinates), and W^k (the spatial components of the four-velocity). For non-rotating dust and a hypersurface-orthogonal foliation, one finds $W_k = 0$. The action of the dust can then be written in the form

$$S^{\mathrm{d}} = \int \mathrm{d}t \int_0^\infty \mathrm{d}r \left(P_\tau \dot{\tau} - N\mathcal{H}^{\mathrm{d}} - N^r \mathcal{H}_r^{\mathrm{d}} \right) \,, \tag{7.120}$$

where the contributions to the Hamiltonian and momentum constraints turn out to read

$$\mathcal{H}^{\mathrm{d}} = P_\tau \sqrt{1 + \frac{{\tau'}^2}{L^2}} \quad \text{and} \quad \mathcal{H}_r^{\mathrm{d}} = \tau' P_\tau \,. \tag{7.121}$$

In order to derive the fall-off conditions analogously to the vacuum case discussed in Section 7.2, one expresses $F(\rho)$, introduced in (7.115), as a function of the canonical variables. The result is

$$F = R\left[1 + G^2\frac{P_L^2}{R^2} - \frac{R'^2}{L^2}\right] . \tag{7.122}$$

As in the Schwarzschild case, the mass function F is elevated to a canonical coordinate by a canonical transformation:

$$(L, P_L; R, P_R; \tau, P_\tau) \longrightarrow (F, P_F; R, \bar{P}_R; \tau, P_\tau), \tag{7.123}$$

where

$$\bar{P}_R = P_R - \frac{LP_L}{2R} - \frac{LP_L}{2R\mathcal{F}} - \frac{\Delta}{RL^2\mathcal{F}} , \tag{7.124}$$

with

$$\Delta = (RR')(LP_L)' - (RR')'(LP_L) \tag{7.125}$$

and

$$\mathcal{F} := 1 - F/R. \tag{7.126}$$

The action in terms of the new canonical variables is then

$$S = \int \mathrm{d}t \int_0^\infty \mathrm{d}r \left(P_\tau\dot{\tau} + \bar{P}_R\dot{R} + P_F\dot{F} - N\mathcal{H} - N^r\mathcal{H}_r\right) + S_{\partial\Sigma} , \tag{7.127}$$

where the new constraints are

$$\mathcal{H} = -\frac{1}{2L}\left(\frac{F'R'}{G\mathcal{F}} + 4G\mathcal{F}P_F\bar{P}_R\right) + P_\tau\sqrt{1 + \frac{\tau'^2}{L^2}} , \tag{7.128}$$

$$\mathcal{H}_r = \tau' P_\tau + R'\bar{P}_R + F'P_F . \tag{7.129}$$

What about the boundary action $S_{\partial\Sigma}$? As was shown in detail in Kiefer *et al.* (2006), it can be absorbed in the course of a canonical transformation that employs the mass density $\Gamma := F'$ as a new canonical variable. One then arrives at a new action

$$S = \int \mathrm{d}t \int_0^\infty \mathrm{d}r \left(P_\tau\dot{\tau} + \bar{P}_R\dot{R} + P_\Gamma\dot{\Gamma} - N\mathcal{H} - N^r\mathcal{H}_r\right) , \tag{7.130}$$

where the constraints in terms of the new variables read

$$\mathcal{H} = -\frac{1}{2L}\left(\frac{\Gamma R'}{G\mathcal{F}} - 4G\mathcal{F}P_\Gamma'\bar{P}_R\right) + P_\tau\sqrt{1 + \frac{\tau'^2}{L^2}} , \tag{7.131}$$

$$\mathcal{H}_r = \tau' P_\tau + R'\bar{P}_R - \Gamma P_\Gamma' . \tag{7.132}$$

The Hamiltonian constraint can be greatly simplified if the momentum constraint is used to eliminate P_F. The constraints (7.131) and (7.132) can then be replaced by the equivalent set

$$H = G\left({P_\tau}^2 + \mathcal{F}\bar{P}_R^2\right) - \frac{\Gamma^2}{4G\mathcal{F}} \approx 0 , \tag{7.133}$$

$$H_r = \tau' P_\tau + R'\bar{P}_R - \Gamma {P_\Gamma}' \approx 0 . \tag{7.134}$$

These equations will be used as the starting point for the quantization discussed below.

We mention that the functions τ_0 and E can also be expressed in terms of the canonical variables, as we have already done for the variable F; see (7.122). This facilitates the interpretation of the canonical variables.

The final form (7.133) of the Hamiltonian constraint was obtained by a series of manipulations. As a consequence, it has acquired some new properties compared with the original form (7.28). First, it possesses the physical dimensions (mass/length)^2. Second, it is no longer hyperbolic: it is hyperbolic only inside the horizon ($\mathcal{F} < 0$), and elliptic outside it ($\mathcal{F} > 0$).

7.5.2 Quantization

We shall now apply the Dirac quantization procedure and turn the classical constraints (7.133) and (7.134) into quantum operators. The translation of Poisson brackets into commutators is achieved in the Schrödinger representation by the formal substitution

$$P_\tau(r) \to \frac{\hbar}{\mathrm{i}}\frac{\delta}{\delta\tau(r)}\,, \quad \bar{P}_R(r) \to \frac{\hbar}{\mathrm{i}}\frac{\delta}{\delta R(r)}\,, \quad P_\Gamma(r) \to \frac{\hbar}{\mathrm{i}}\frac{\delta}{\delta\Gamma(r)} \tag{7.135}$$

and acting with them on wave functionals. The Hamiltonian constraint (7.133) then leads to the Wheeler–DeWitt equation,

$$\begin{aligned}\Bigg[-G\hbar^2\bigg(\frac{\delta^2}{\delta\tau(r)^2} + \mathcal{F}\frac{\delta^2}{\delta R(r)^2} + A(R,F)\,\delta(0)\,\frac{\delta}{\delta R(r)} \\ + B(R,F)\,\delta(0)^2\bigg) - \frac{\Gamma^2}{4G\mathcal{F}}\Bigg]\Psi\left[\tau(r'), R(r'), \Gamma(r')\right] = 0\,,\end{aligned} \tag{7.136}$$

where A and B are smooth functions of R and F that encapsulate the factor-ordering ambiguities. We have introduced divergent quantities such as $\delta(0)$ in order to indicate that the factor-ordering problem is unsolved at this stage and can be dealt with only after some suitable regularization has been performed (see Isham 1976). That is, in the ideal case one would like to choose the terms proportional to $\delta(0)$ in such a way that the constraint algebra closes; cf. Section 5.3.5.

Quantizing the momentum constraint (7.134) by means of (7.135), we obtain

$$\left[\tau'\frac{\delta}{\delta\tau(r)} + R'\frac{\delta}{\delta R(r)} - \Gamma\left(\frac{\delta}{\delta\Gamma(r)}\right)'\right]\Psi\left[\tau(r'), R(r'), \Gamma(r')\right] = 0\,. \tag{7.137}$$

Up to now, the quantum constraint equations have been given only in a formal way. We shall now attempt to define them by a lattice regularization.

We consider a one-dimensional lattice given by a discrete set of points r_i separated by a distance σ. To fulfill the momentum constraint in the continuum limit, it is important to start with a corresponding ansatz for the wave functional before putting it on the lattice. We choose the form

$$\Psi\left[\tau(r),\, R(r),\, \Gamma(r)\right] = U\left(\int \mathrm{d}r\,\Gamma(r)\mathcal{W}(\tau(r),\, R(r),\, \Gamma(r))\right)\,, \tag{7.138}$$

where $U : \mathbb{R} \to \mathbb{C}$ is at this stage some arbitrary (differentiable) function. Using Γ in the exponent instead of R' or τ' is suggested by the form of the Wheeler–DeWitt equation (absence of derivatives with respect to Γ) and the fact that $F' = \Gamma$ is related to the energy density. The ansatz has to be compatible with the lattice, which means that it has to factorize into different functions for each lattice point. So we have to make the choice $U = \exp$, which gives

$$\begin{aligned}
&\Psi\left[\tau(r),\, R(r),\, \Gamma(r)\right] \\
&= \exp\left(\int \mathrm{d}r\, \Gamma(r)\mathcal{W}(\tau(r),\, R(r),\, F(r))\right) \\
&= \exp\left(\lim_{\sigma\to 0}\sum_i \sigma\Gamma_i\mathcal{W}_i\left(\tau(r_i),\, R(r_i),\, F(r_i)\right)\right) \\
&= \lim_{\sigma\to 0}\prod_i \exp\left(\sigma\,\Gamma_i\,\mathcal{W}_i\left(\tau(r_i),\, R(r_i),\, F(r_i)\right)\right) \\
&\equiv \lim_{\sigma\to 0}\prod_i \Psi_i\left(\tau(r_i),\, R(r_i),\, \Gamma(r_i),\, F(r_i)\right)\ ,
\end{aligned} \tag{7.139}$$

where

$$F(r_i) = \sum_{j=0}^{i} \sigma\,\Gamma_j\ . \tag{7.140}$$

We implement the formal expression $\delta(0)$ on the lattice as

$$\delta(0) \to \lim_{\sigma\to 0}\frac{1}{\sigma}\ . \tag{7.141}$$

The lattice version of the Wheeler–DeWitt equation (7.136) then reads

$$\left[G\hbar^2\left(\frac{\partial^2}{\partial\tau_j^2} + \mathcal{F}_j\,\frac{\partial^2}{\partial R_j^2} + A(R_j, F_j)\,\frac{\partial}{\partial R_j}\right) + B(R,F) + \frac{\sigma^2\,\Gamma}{4G\,\mathcal{F}_j}\right]\Psi_j = 0\ . \tag{7.142}$$

We now insert the ansatz (7.139) and make for convenience the redefinition $\mathcal{W} = \mathrm{i}W/2$. This leads to

$$\begin{aligned}
&\frac{\sigma^2\Gamma_i^2}{4}\left[G\hbar^2\left(\frac{\partial W(\tau,\, R,\, F)}{\partial\tau}\right)^2 + G\hbar^2\mathcal{F}\left(\frac{\partial W(\tau,\, R,\, F)}{\partial R}\right)^2 - \frac{1}{G\mathcal{F}}\right] \\
&\qquad + \frac{\sigma\Gamma_i}{2}\left[G\hbar^2\left(\frac{\partial^2}{\partial\tau^2} + \mathcal{F}\,\frac{\partial}{\partial R^2} + A(R,F)\frac{\partial}{\partial R}\right)W(\tau,\, R,\, F)\right] \\
&\qquad\qquad + B(R,F) = 0\ .
\end{aligned} \tag{7.143}$$

In order for this to be fulfilled independently of the choice of σ (and thus also in the limit $\sigma \to 0$), one is led to the following *three* equations:

$$\left(G\hbar\frac{\partial W(\tau,\, R,\, \Gamma)}{\partial\tau}\right)^2 + \mathcal{F}\left(G\hbar\frac{\partial W(\tau,\, R,\, \Gamma)}{\partial R}\right)^2 - \frac{1}{\mathcal{F}} = 0\ , \tag{7.144}$$

$$\left(\frac{\partial^2}{\partial\tau^2} + \mathcal{F}\,\frac{\partial^2}{\partial R^2} + A(R,\Gamma)\frac{\partial}{\partial R}\right)W(\tau,\, R,\, \Gamma) = 0\ , \tag{7.145}$$

and

$$B(R, \Gamma) = 0. \tag{7.146}$$

The first equation (7.144) is the Hamilton–Jacobi equation. The second equation presents an additional restriction on solutions of (7.144). The last equation (7.146) tells us that working on the lattice is only possible if the factor ordering does not contribute to the potential term. If we can find solutions to all three equations, we can do all other calculations on the lattice, since these solutions have a well-defined continuum limit and satisfy the momentum constraint.

At this point, we have to make a few comments on the regularization procedure we are using. It has already been noted that the lattice regularization does not solve the factor-ordering problem. The lattice regularization is simply an ad hoc regularization in which the divergent terms have to cancel each other. Put differently, it is equivalent to a type of regularization that was employed in DeWitt (1967*a*) (which means setting $\delta(0) = 0$) with an additional constraint on the solutions.[8]

We have already mentioned, at the end of the last subsection, that the signature in the kinetic part of the Hamiltonian constraint (7.133) changes from elliptic (outside the horizon) to hyperbolic (inside the horizon). This thus occurs for the kinetic term of the Wheeler–DeWitt equation (7.136), too. As discussed in Brotz and Kiefer (1997), we can say that the part inside the horizon is always classically allowed, whereas this is not necessarily the case for the outside part. The usual initial-value problem appropriate to hyperbolic equations can thus only be applied to the region corresponding to the black-hole interior.

The two equations (7.144) and (7.145) have to be satisfied in order to get a diffeomorphism-invariant solution to the Wheeler–DeWitt equation. Looking for particular solutions of the separating form $W = \alpha(\tau) + \beta(R)$, we recognize immediately that this system of equations is consistent *only for special factor orderings*. Tackling the problem from the opposite point of view, one can ask which factor orderings *do* yield a separating solution. We find

$$A(R; F) = \frac{F}{2R^2} \frac{2 - a^2\mathcal{F}}{1 - a^2\mathcal{F}} \, . \tag{7.147}$$

This leads to

$$G\hbar W(\tau, \Gamma, R, a) = \text{const.} \pm a\tau \pm \int dR \, \frac{\sqrt{1 - a^2\mathcal{F}}}{\mathcal{F}} \, . \tag{7.148}$$

It turns out that the functions W are also solutions to the Hamilton–Jacobi equation. From this, one gets the identification $2E = 1/a^2 - 1$. Since classically $E > -1/2$, it follows that a should be real.

Recalling that $\mathcal{W} = \mathrm{i}W/2$, we recognize from (7.139) that the wave functionals in the continuum limit read (up to a normalization factor)

$$\Psi\left[\tau(r), R(r), \Gamma(r)\right] = \exp\left(\pm\frac{\mathrm{i}}{2G\hbar}\int \mathrm{d}r\ \Gamma\left[a\tau + \int^R \mathrm{d}R\ \frac{\sqrt{1 - a^2\mathcal{F}}}{\mathcal{F}}\right]\right) . \tag{7.149}$$

[8] The condition $\delta(0) = 0$ is also used in dimensional regularization; see e.g. (1.98) in Hamber (2009).

The integral appearing in (7.148) can be evaluated exactly. One gets

$$\int \mathrm{d}R\, \frac{\sqrt{1-a^2\mathcal{F}}}{\mathcal{F}}$$
$$= F\left[\frac{\sqrt{1-a^2\mathcal{F}}}{1-\mathcal{F}} - 2\tanh^{-1}\sqrt{1-a^2\mathcal{F}} + \frac{2-a^2}{\sqrt{1-a^2}}\tanh^{-1}\frac{\sqrt{1-a^2}}{\sqrt{1-a^2\mathcal{F}}}\right]. \quad (7.150)$$

Here we have assumed that $0 < a \leq 1$ (corresponding to $E \geq 0$). For $a > 1$ ($E < 0$) but $R < Fa^2/(a^2-1)$, the result can be found by analytic continuation to read

$$F\left[\frac{\sqrt{1-a^2\mathcal{F}}}{1-\mathcal{F}} + \ln\left|\frac{1-\sqrt{1-a^2\mathcal{F}}}{1-\sqrt{1-a^2\mathcal{F}}}\right| + \frac{2-a^2}{\sqrt{a^2-1}}\tan^{-1}\frac{\sqrt{a^2-1}}{\sqrt{1-a^2\mathcal{F}}}\right]. \quad (7.151)$$

Another analytic continuation gives the result in the region $R > Fa^2/(a^2-1)$:

$$\mathrm{i}F\left[\frac{\sqrt{a^2\mathcal{F}-1}}{1-\mathcal{F}} - 2\tan^{-1}\sqrt{a^2\mathcal{F}-1} + \frac{(1-a^2/2)}{\sqrt{a^2-1}}\ln\left|\frac{\sqrt{a^2\mathcal{F}-1}-\sqrt{a^2-1}}{\sqrt{a^2\mathcal{F}-1}+\sqrt{a^2-1}}\right|\right]. \quad (7.152)$$

We recognize from (7.152) that the wave function becomes a real exponential in the region $R > Fa^2/(a^2-1) = -F/2E$. This can be interpreted by rewriting the classical Einstein equations in the form

$$(\partial_\tau R)^2 = 2E + 1 - \mathcal{F} = 2E + \frac{F}{R}. \quad (7.153)$$

Therefore, the wave function is real in the region that is classically forbidden.

Surprisingly, (7.148) in fact gives the *complete* class of solutions with the form (7.139); see Kiefer *et al.* (2006). There thus exist no non-separating solutions under the assumption that the full wave functional factorizes into functions on the respective lattice points. But we emphasize that we have succeeded in finding exact solutions to all quantum constraints. Other solutions to the full Wheeler–DeWitt equation and momentum constraints would necessarily couple the infinitely many shells comprising the dust cloud; to find them would demand a regularization scheme that is much more involved. In a sense, the factor ordering chosen here leads to quantum states for which the WKB form is 'exact'; compare the analogous situation for the models discussed in Brotz and Kiefer (1997) and Louis-Martinez *et al.* (1994). Alternatively, one can relax the demand for an exact solution and make a Born–Oppenheimer type of ansatz in the spirit of Section 5.4 (Banerjee *et al.* 2010). One thereby arrives at the same solutions (7.149) at order l_P^2. The higher orders then lead to quantum-gravitational corrections. These corrections are absent in the above treatment because of the special 'fine-tuned' choice (7.147) for the factor ordering.

Most of the above expressions retain their form in the presence of a cosmological constant Λ (Franzen *et al.* 2010). Instead of (7.126), one now has

$$\mathcal{F} := 1 - \frac{F}{R} - \frac{\Lambda R^2}{3}.$$

The constraints and the solutions assume the same form with the new $\mathcal{F}$. The only exception is the factor-ordering term A, which instead of (7.147) reads

$$A = \frac{1}{2}\left(\frac{F}{R^2} - \frac{2\Lambda R}{3}\right)\frac{2 - a^2\mathcal{F}}{1 - a^2\mathcal{F}}$$

here. Outside the dust cloud, the corresponding classical solution is the Schwarzschild–de Sitter (or Kottler) solution.

What has been achieved? It was possible to solve the coupled Wheeler–DeWitt equation and momentum constraint, at least for a special factor ordering. This is already a non-trivial result. Alternatively, one can find a solution in the semiclassical approximation. However, one can also see the limits. An interesting application would be to the fate of the classical singularities (both black-hole and naked singularities) in the quantum theory. Such a discussion was possible for the thin shell of Section 7.4. But in the present case one would have to go beyond the special set of solutions (7.149), which has not yet been achieved. Some attempts at studying singularity avoidance have, however, been made in the loop-quantum-gravity version of the LTB model (Bojowald *et al.* 2009).

Applications have thus so far been restricted to the semiclassical states (7.149). These are concerned with Hawking radiation and the statistical interpretation of black-hole entropy. Let us turn first to Hawking radiation (Kiefer *et al.* 2007*b*). The Wheeler–DeWitt equation (7.136) contains second derivatives with respect to the dust time τ. We can thus define states of positive and negative frequency with respect to τ, depending on the sign of the imaginary unit in (7.149). This situation is analogous to the Klein–Gordon equation. Hawking radiation is then found from the overlap of an 'outgoing dust state' with negative frequency and an 'ingoing dust state' with positive energy. The result is the standard expression $k_{\mathrm{B}}T_{\mathrm{BH}} = \hbar/8\pi GM$ for the Hawking temperature of a spherically symmetric black hole. The calculation can also be performed for the Schwarzschild–de Sitter solution, and the two standard expressions for the Hawking temperature (corresponding to the black-hole and the cosmological horizon, respectively) are recovered too (Franzen *et al.* 2010). The same holds for a black hole in AdS space (Vaz *et al.* 2008*b*).

What about the entropy? In a heuristic attempt, the black hole has been interpreted as a single shell that arises as the end state of the collapse of the many shells comprising the dust cloud in the LTB model (Vaz *et al.* 2008*a*). It is assumed that each shell can occupy only the energy levels calculated from the Wheeler–DeWitt equation (7.136). Following the ideas of Bekenstein (1973), the black-hole entropy is then interpreted as the number of possible distributions of a given number of shells between these levels. The desired result $A/4G\hbar$ for the entropy is then obtained if an integration constant appearing in this calculation is chosen appropriately. This provides an alternative approach to developing a statistical foundation for black-hole entropy, which is different from the approaches based on loop quantum gravity and string theory.

7.6 The information-loss problem

What happens during the final evaporation phase of a black hole? This is, of course, one of the major questions that a quantum theory of gravity should answer. So far, only simplified models exist; see, for example, Kiefer *et al.* (2009). As long as this question remains open, various gedanken experiments and heuristic expectations will be discussed. One of these concerns the so-called information-loss problem; see Page (1994) or Mathur (2011) and the references therein for a detailed exposition. According to the semiclassical calculations that lead to (7.11), black-hole radiation has a thermal spectrum. As a consequence, the black hole loses mass and shrinks. If the hole evaporates completely and leaves only thermal radiation behind, *any* initial state for the black hole plus the quantum field will end up in the *same* final state, which will be a thermal, that is, mixed, state. This corresponds to a maximal loss of information about the initial state. In other words, unitarity would be violated even for a closed system, in contrast to standard quantum theory. Formally, the expression $\operatorname{tr}\rho^2$ remains *constant* under the von Neumann equation. The same is true for the entropy $S_{\rm SM} = -k_{\rm B}\operatorname{tr}(\rho\ln\rho)$: for a unitarily evolving system, there is no increase in entropy. If these laws were violated during black-hole evaporation, information would be destroyed. The attitudes towards this information-loss problem can be roughly divided into the following classes:

1. The information is indeed lost during black-hole evaporation, and the quantum-mechanical Liouville equation is replaced by an equation of the form

$$\rho \longrightarrow \$\rho \neq S\rho S^\dagger \,, \tag{7.154}$$

   where \$ is Hawking's 'dollar matrix', which generalizes the ordinary S-matrix S. Hawking originally assumed that the exact evolution was non-unitary and that information was lost (Hawking 1976), but he later changed his mind (Hawking 2005).
2. The full evolution is in fact unitary; the black-hole radiation contains subtle quantum correlations that cannot be seen in the semiclassical approximation.
3. The black hole does not evaporate completely but leaves a 'remnant' with a mass of the order of the Planck mass, which carries the whole information.

We shall not address the last option. A concrete realization of this possibility is provided by the asympotic-safety approach discussed in Section 2.2.5; see Bonanno and Reuter (2006).

As we discussed at length in Chapter 5, in full quantum gravity there is no notion of external time. However, for an isolated system such as a black hole, one can refer to the semiclassical time of external observers far away from the hole. The notion of unitarity then refers to this concept of time (Kiefer *et al.* 2009). Consequently, if the fundamental theory of quantum gravity is unitary in this sense, there will be no information loss. Conversely, if the fundamental theory breaks unitarity in this sense, information loss is possible. As long as the situation about the full theory remains open, discussions of the information loss must centre around assumptions and expectations (Mathur 2011).

For a black hole whose mass is much greater than the Planck mass, the details of quantum gravity should be less relevant. It is of importance to emphasize that a (large) black hole is a macroscopic object. It is therefore strongly entangled with the quantum degrees of freedom with which it interacts; the black hole is an *open* quantum system. Information thereby becomes essentially non-local. Because of this entanglement, the black hole can assume classical properties like other macroscopic objects. This emergence of classical properties through interaction with other degrees of freedom is called *decoherence*; cf. Section 10.2. A particular consequence is that the black hole itself cannot evolve unitarily, only the total system consisting of the black hole and interacting fields (Zeh 2005). As was shown in Demers and Kiefer (1996), the interaction of the black hole with its Hawking radiation is sufficient to provide the black hole with classical behaviour; strictly speaking, the very notion of a black hole emerges through decoherence. The cases of a superposition of a black-hole state with its time-reversed version (a 'white hole') and with a no-hole state were also considered and shown to decohere by Hawking radiation. This does not hold for virtual black holes, which are time-symmetric and do not exhibit classical behaviour.

In the original derivation of Hawking radiation, one starts with a quantum field that is in its vacuum state; cf. Hawking (1975) and Section 7.1.2. The thermal appearance of the resulting Hawking radiation is then recognized from the Planckian form of the expectation value of the particle-number operator at late times. If the state of the quantum field (which is a two-mode squeezed state) is evaluated on a spatial hypersurface that enters the horizon (but without encountering the singularity), tracing out the degrees of freedom of the hole interior yields a thermal density matrix with the temperature given in (7.11) in the outside region; cf. Section 7.1.2. One is not, however, obliged to take a hypersurface that enters the horizon. One can consider a hypersurface that is locked at the bifurcation point (where the surface of the collapsing star crosses the horizon). Then, apart from its entanglement with the black hole itself, the field state remains pure. Observations far away from the hole should not, however, depend on the choice of the hypersurface (Zeh 2005). This is, in fact, what results (Kiefer 2001*b*, 2004*a*): the entanglement of the squeezed state representing the Hawking radiation with other (irrelevant) degrees of freedom leads to the thermal *appearance* of the field state (cf. also Hsu and Reeb (2009)). It therefore seems that there is no information-loss problem at the semiclassical level and that one can assume, at the present level of understanding, that the full evolution is unitary and that 'information loss' can be understood in terms of the standard delocalization of information due to decoherence. There are indications that a similar scenario can emerge from the string-theory approach discussed in Chapter 9.

Quantum entanglement could be the origin of the Bekenstein–Hawking entropy. If some degrees of freedom are traced out in an entangled state, the remaining variables possess a positive entanglement entropy; see, for example, Das *et al.* (2008). But what are the relevant degrees of freedom? This question is difficult to answer, because the resulting entropy should not depend on the particular state and fields, but should yield a *universal* result—the Bekenstein–Hawking entropy (7.23).[9] One suggestion is to use the quasi-normal modes (Section 7.3.2) as the degrees of freedom that are entangled with the microscopic degrees of freedom of the black hole (Kiefer 2004*a*). The reason is the universal nature of these modes. However, no conclusive calculation exists so far.

[9]This puzzle of the universality of the entropy is one of the five problems in quantum gravity presented in Strominger (2009); the remaining four open problems also deal with black holes.

7.7 Primordial black holes

A major test of many of the issues discussed in this chapter would of course be an experimental test of the Hawking effect. As is clear from (7.13), the Hawking temperature is large enough only for small black-hole masses—much smaller than the masses of black holes that result from stellar collapse. To form such black holes, one needs densities that can only occur under the extreme conditions of the early universe. These objects are therefore called *primordial black holes* (PBHs).[10] They can originate in the radiation-dominated phase, during which no stars or other objects can be formed. We shall give a brief summary of this topic here. For more details, see Carr (2003), Carr *et al.* (2010), and the references therein.

In order to study the formation of PBHs, consider for simplicity a spherically symmetric region with radius R and density ρ embedded in a flat universe with critical density $\rho_{\rm c}$. According to the Tolman–Oppenheimer–Volkov (TOV) solution for constant density (incompressible matter), we have

$$M(R) = \frac{4\pi}{3}\rho R^3.$$

In the limit $R \to R_{\rm S} = 2GM/c^2$ (reinserting c for the moment), this would lead to a 'PBH black-hole density'

$$\rho_{\rm PBH} = \frac{M}{4\pi R_{\rm S}^3/3} = \frac{3c^6}{32\pi G^3 M^2} \approx 1.84 \times 10^{16} \left(\frac{M_\odot}{M}\right)^2 \frac{\rm g}{\rm cm^3},$$

which thus increases with decreasing mass. This density has to be compared with the cosmic density

$$\rho = \rho_{\rm c} = \frac{3H^2}{8\pi G}.$$

Assuming a radiation-dominated flat universe, where $H(t) = (2t)^{-1}$, we get from the condition $\rho_{\rm PBH} \approx \rho_{\rm c}$ that

$$M(t) \approx \frac{c^3 t}{G} \approx M_* \left(\frac{t}{10^{-24}\ {\rm s}}\right) \approx 1 M_\odot \left(\frac{t}{5 \times 10^{-6}\ {\rm s}}\right), \tag{7.155}$$

where $M_* \approx 5 \times 10^{14}$ g is the mass of a black hole whose lifetime coincides with the age of our Universe; see (7.20). If we compare this with the particle horizon of a radiation-dominated universe, $d_H(t) = cH^{-1}(t) = 2ct$, we get

$$M(t) \approx M_H(t), \tag{7.156}$$

where $M_H(t) = 4\pi\rho_{\rm c} d_H^3/3$ is the horizon mass. This rough estimate thus indicates that the mass of a primordial black hole, if it is formed, is of the order of the horizon mass. At the Planck time $t_{\rm P}$ we thus expect, not surprisingly, a PBH with the Planck mass $m_{\rm P}$, at 10^{-24} s one with mass M_*, at 5×10^{-6} s one with a solar mass, and at 5 s (the time of nucleosynthesis) one with a mass of a million solar masses, comparable

[10] The possibility of generating small black holes in accelerators is discussed in Section 9.2.7.

to our Galactic Black Hole, which has a mass of about $4 \times 10^6 M_\odot$. The initial mass can increase by means of accretion, but it turns out that this is negligible under most circumstances.

We would expect PBH formation from the Planck time up to a cosmic age of about one second in the radiation-dominated universe. If there was an inflationary state in the early universe (at around 10^{-34} to 10^{-32} s), all PBHs formed during inflation would be diluted away. This would then give the bound (Carr 2003)

$$M_{\mathrm{PBH}} > M_H(T_{\mathrm{RH}}) \approx \frac{m_{\mathrm{P}}^3 c^4}{10.88(k_{\mathrm{B}} T_{\mathrm{RH}})^2} \sim 1 \text{ g}, \tag{7.157}$$

if a value of 10^{16} GeV is chosen for the reheating temperature T_{RH} .

Since (7.156) is only a rough estimate, one introduces a parameter γ that gives the ratio of the PBH mass to the horizon mass (Carr *et al.* 2010),

$$M = \gamma M_H = 2.03 \times 10^5 \gamma \left(\frac{t}{1 \text{ s}} \right) M_\odot .$$

What is γ? If one assumes that the radius R of the collapsing region must be bigger than the Jeans length

$$\lambda_{\mathrm{J}} := \sqrt{\frac{1}{3}} d_H ,$$

one has $\gamma = (1/\sqrt{3})^3 \approx 0.2$.[11] However, in more precise calculations, somewhat different values are obtained, which is why γ is left as a free parameter (Carr *et al.* 2010). We note that R must be smaller than the curvature radius of the overdense region at the moment of collapse, which is roughly equal to H^{-1}. Otherwise, the region would contain a compact three-sphere that was topologically disconnected from the rest of the universe; this case would not lead to a black hole within *our* universe. One therefore has the condition

$$H^{-1} \gtrsim R \gtrsim \sqrt{\frac{1}{3}} H^{-1}, \tag{7.158}$$

evaluated at the time of collapse, for the formation of a black hole. This relation can also be rewritten as a condition referring to any initial time of interest. In particular, one is often interested in the time at which the fluctuation enters the Hubble scale in the radiation-dominated universe. This is illustrated in Fig. 7.6, where the presence of a possible inflationary phase at earlier times (see below) is also shown.

PBHs with an initial mass M_* are distinguished because they would evaporate today. (They would release an energy of about 10^{30} erg in the last second.) They would thus be most important for quantum gravity, because the final evaporation phase is one of the key puzzles that such a theory is supposed to solve. It follows from (7.13) that PBHs with a mass M_* would produce γ-rays of order 100 keV and would thus contribute to the observed γ-ray background in a distinguished way given by the thermal nature of the evaporation. Since no such contribution is observed, Page and

[11] The term $1/\sqrt{3}$ comes from the relation $p = w\rho$, with $w = 1/3$ for radiation.

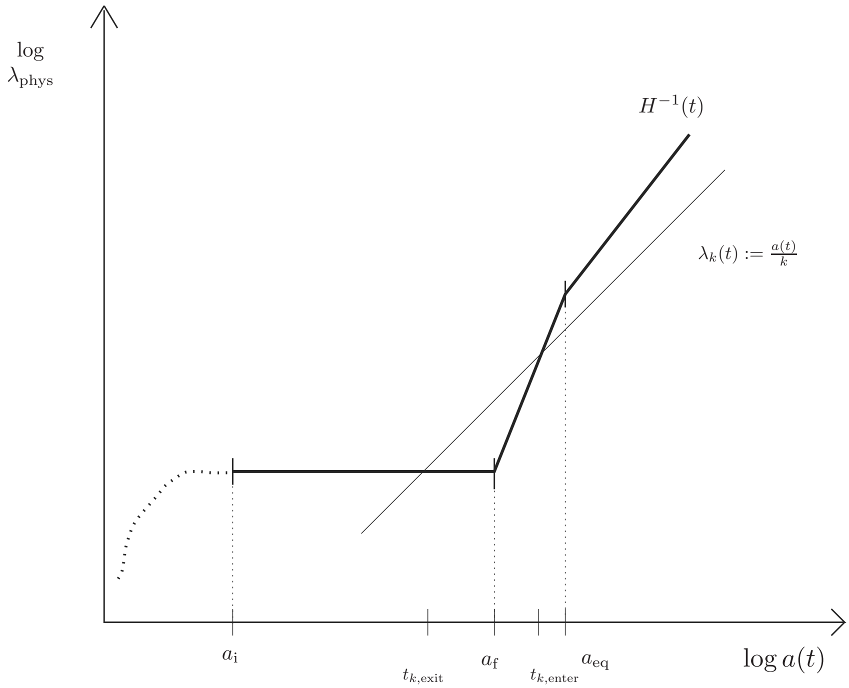

Fig. 7.6 Time development of a physical scale $\lambda(t) := a(t)/k$, where $a(t)$ is the scale factor of a Friedmann universe and k is the dimensionless wave number, and the Hubble scale $H^{-1}(t)$. During an inflationary phase, $H^{-1}(t)$ remains approximately constant. After the end of inflation (a_{f}), the Hubble scale $H^{-1}(t)$ increases faster than any physical scale. Therefore the scale described by λ_k, which has left the Hubble scale at time $t_{k,\text{exit}}$, enters the Hubble scale again at $t_{k,\text{enter}}$ in the radiation- (or matter-)dominated phase.

Hawking (1976) were able to derive the following upper limit on the contribution of such PBHs to the present energy content of the Universe:

$$\Omega_{\text{PBH}}(M_*) := \frac{\rho_{\text{PBH}}(M_*)}{\rho_{\text{c}}} < 10^{-8}.$$

If these black holes are uniformly distributed, this leads to an upper limit of 10^{-6} explosions per cubic parsec per year. (If they are clustered inside galactic halos, weaker constraints are obtained.)[12]

PBHs with masses greater than M_* would still exist today and could even contribute in significant amounts to the cold dark matter (CDM). We shall discuss an example below. PBHs with masses smaller than M_* would have completely evaporated by now. However, one could detect their previous existence by looking at processes influenced by the evaporation. Carr *et al.* (2010) have investigated many such processes

[12]There are, however, claims that some short-period γ-ray bursts may be due to PBH explosions (Cline *et al.* 2003).

and have found, in particular, a striking modification of the big bang nucleosynthesis scenario; from the success of the unmodified scenario, one can then find strong limits on the corresponding PBH abundance (see Fig. 7.7 below).

An important parameter in the study of PBH formation is the mass fraction, $\beta(M)$, of the universe that goes into PBHs at the time $t_{\rm i}$ of formation:

$$\beta(M) = \frac{\rho_{\rm PBH}(t_{\rm i})}{\rho(t_{\rm i})}. \tag{7.159}$$

It is usually assumed that this mass function is monochromatic; that is, at a given $t_{\rm i}$, all PBHs have the same mass. Strictly speaking, this is not correct; according to numerical calculations by Niemeyer and Jedamzik (1999), there exists a whole spectrum of initial masses,

$$M_{\rm PBH} = KM_H(\delta - \delta_{\rm min})^{\xi}, \tag{7.160}$$

a relation that is reminiscent of the theory of critical phenomena. This may change some of the quantitative conclusions.

An immediate upper limit on $\beta(M)$ comes from the requirement that the PBH mass density today cannot exceed the observational limit on the CDM density. Setting $\Omega_{\rm PBH} < \Omega_{\rm CDM} < 0.25$, Carr *et al.* (2010) present the constraint

$$\beta(M) < 2.04 \times 10^{-18}\gamma^{-1/2}\left(\frac{\Omega_{\rm CDM}}{0.25}\right)\left(\frac{h}{0.72}\right)^2\left(\frac{g_{*\rm i}}{106.75}\right)^{1/4}\left(\frac{M}{10^{15}\ {\rm g}}\right)^{1/2},$$

which is valid for $M > M_*$. In this expression, h is the dimensionless Hubble parameter (current observations indicate $h \approx 0.72$) and $g_{*\rm i}$ is the number of relativistic degrees of freedom at the time of PBH formation. To get this relation, the transition from the radiation- to the matter-dominated phase must be taken into account properly. Carr *et al.* (2010) introduce for convenience the new parameter

$$\beta'(M) := \gamma^{1/2}\left(\frac{g_{*\rm i}}{106.75}\right)^{-1/4}\beta(M). \tag{7.161}$$

Figure 7.7 collects together the empirically known constraints on $\beta'(M)$ for the mass range 1–10^{50} g. It is evident that these constraints are already tight; that is, the small numbers for the upper limits of $\beta'(M)$ indicate that the PBH production rate was very low. Not surprisingly, the strongest constraints on $\beta'(M)$ occur in the mass range around M_*. Among the missions that are searching for PBH signatures is the Fermi Gamma-ray Space Telescope, launched in 2008, which is dedicated to performing γ-ray astronomy observations.

PBH production can be investigated in particular in the context of inflationary models for the early universe; see, for example, Blais *et al.* (2003). Inflation can produce a spectrum of fluctuations that is relevant to PBH formation. For this, one needs to determine the size of the fluctuations when they re-enter the Hubble scale in the radiation-dominated phase, that is, when their wavelength becomes smaller than $H^{-1} = \dot{a}/a$; see Fig. 7.6. For the calculation, one needs, of course, an initial spectrum for the fluctuations. This is usually taken to be of a Gaussian form, as predicted by

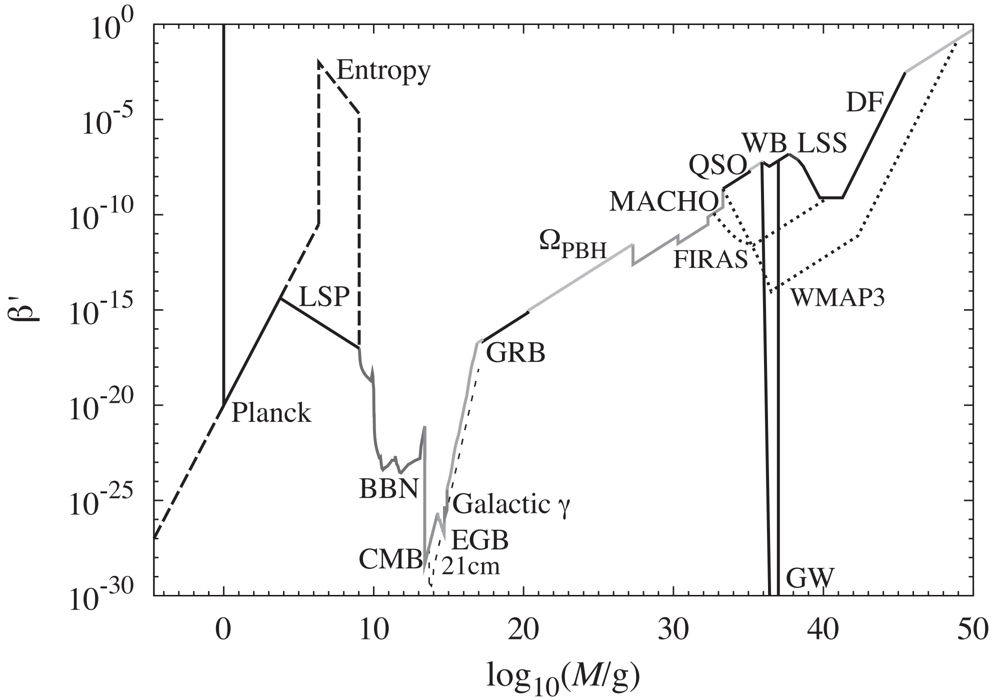

Fig. 7.7 Constraints on $\beta'(M)$ in the mass range 1–10^{50} g. The acronyms are as follows: gamma-ray bursts (GRB), microlensing of stars (MACHO), quasars (QSO), wide binary disruption (WB), dynamical friction (DF), generation of large-scale structure through Poisson fluctuations (LSS), accretion effects on the CMB (FIRAS, WMAP3), gravitational-wave limits (GW), big bang nucleosynthesis (BBN), extragalactic background (EGB), and lightest supersymmetric particle (LSP). Reprinted with kind permission from Carr *et al.* (2010). © 2010 by the American Physical Society.

most inflationary models (cf. Liddle and Lyth 2000). Therefore, although the fluctuations are small on average, there is always a non-vanishing probability that the density contrast may be high enough to form a black hole.

Assuming that inflation leads to a scale-free power spectrum for the primordial curvature perturbation,

$$k^3 P(k) \propto \left(\frac{k}{k_0}\right)^{n_s - 1},$$

a significant number of PBHs can only be produced for a 'blue spectrum', that is, for a spectrum with $n_s > 1$, where n_s is the spectral index. Only then does one have enough power at the small scales (large k) where PBH formation occurs. Present observations (WMAP-7 data), however, give a value

$$n_s = 0.968 \pm 0.012 \quad (68\% \text{ C.L.}),$$

so this possibility is observationally excluded. Some viable options are models with broken scale invariance (BSI). Let us briefly review such a model to see how this works (Blais *et al.* 2003).

Assuming that the primordial fluctuations obey Gaussian statistics, the probability for the density contrast δ reads

$$p(\delta) = \frac{1}{\sqrt{2\pi}\ \sigma(R)}\ \exp\left(-\frac{\delta^2}{2\sigma^2(R)}\right),$$

where the dispersion (mass variance) can be calculated from the power spectrum $P(k)$:

$$\sigma^2(R) = \frac{1}{2\pi^2}\int_0^\infty dk\ k^2\ W_{\mathrm{TH}}^2(kR)P(k),$$

where $W_{\mathrm{TH}}(kR)$ is the 'top-hat window function' which is needed to smear out small scales in an appropriate way. One then gets the following for the probability that a PBH is formed with mass $M_{\mathrm{PBH}} \geq M_H(t_k)$, where t_k denotes the scale-dependent time at which the fluctuations re-enter the Hubble scale:

$$\beta(M_H) = \frac{1}{\sqrt{2\pi}\,\sigma_H(t_k)}\int_{\delta_{\min}}^{\delta_{\max}} \exp\left(-\frac{\delta^2}{2\sigma_H^2(t_k)}\right)\mathrm{d}\delta \approx \frac{\sigma_H(t_k)}{\sqrt{2\pi}\,\delta_{\min}}\exp\left(-\frac{\delta_{\min}^2}{2\sigma_H^2(t_k)}\right).$$

Numerical calculations indicate that $\delta_{\min} \approx 0.7$ (Niemeyer and Jedamzik 1999).

Let us consider an inflationary model with a jump in the first derivative of the inflaton potential $V(\phi)$ at some scale k_s, with a corresponding value ϕ_s of the inflaton field. This model was discussed by Starobinsky (1992), who presented the following analytic form for the ensuing power spectrum:

$$\frac{9}{4}\left(\frac{aH}{k}\right)^4 k^3 P(k,t_k) = \frac{2H_s^6}{A_-^2}\left[1 - 3(p-1)\frac{1}{y}\left(\left(1-\frac{1}{y^2}\right)\sin 2y + \frac{2}{y}\cos 2y\right)\right.$$
$$\left. + \frac{9}{2}(p-1)^2\frac{1}{y^2}\left(1+\frac{1}{y^2}\right)\left(1+\frac{1}{y^2}+\left(1-\frac{1}{y^2}\right)\cos 2y - \frac{2}{y}\sin 2y\right)\right],$$

where

$$y := \frac{k}{k_s}, \qquad p := \frac{A_-}{A_+}, \qquad H_s^2 = \frac{8\pi G V(\phi_s)}{3},$$

and where $A_-,\ A_+$ are the derivatives of the inflaton potential on the two sides of the jump. The model can thus be characterized by the location k_s of the jump and the parameter p.

A numerical evaluation of the mass fraction $\beta(M_H)$ shows the existence of a peak at a mass M_{peak} near the value $M_s := M_H(t_{k_s}$. By demanding that the present PBH density parameter obeys

$$\Omega_{\mathrm{PBH},0} \approx \Omega_{\mathrm{PBH},0}(M_{\mathrm{peak}}) \leq 0.3,$$

one can calculate the allowed range in the parameter space given by p and M_s (Blais *et al.* 2003). This is shown in Fig. 7.8.

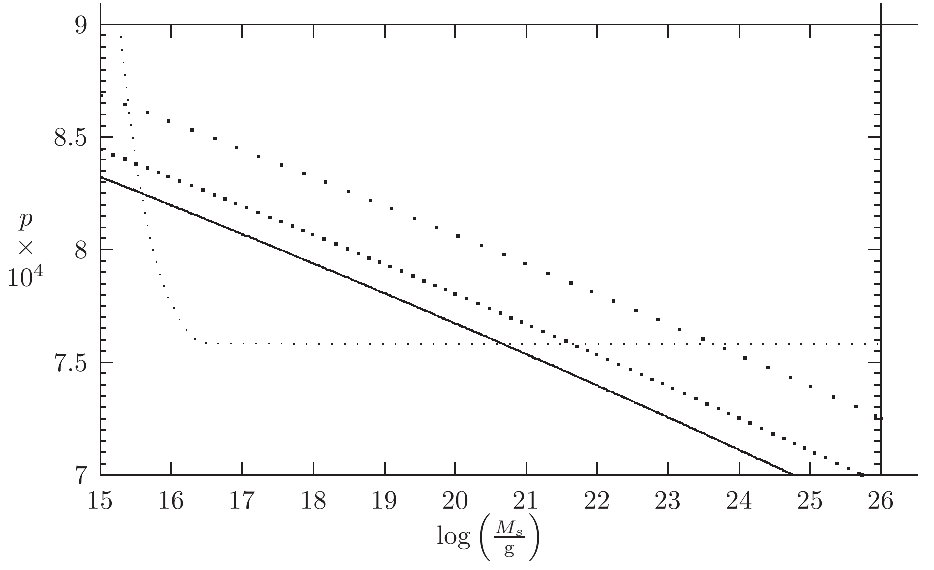

Fig. 7.8 The allowed region in parameter space $(p, M_s = M_H(t_{k_s}))$. The solid line represents those points for which $\Omega_{\mathrm{PBH},0}(M_{\mathrm{peak}}) = 0.3$. Below the solid line, we have $\Omega_{\mathrm{PBH},0}(M_{\mathrm{peak}}) > 0.3$. The two lines parallel to the solid line represent, from bottom to top, those points for which $\Omega_{\mathrm{PBH},0}(M_{\mathrm{peak}}) = 0.1$ and 0.01, respectively. Below the curved dotted line, the γ-ray background constraint is violated. From Blais *et al.* (2003). © 2003 by the American Physical Society.

One recognizes that, in the context of this model, a significant part of cold dark matter could exist in the form of primordial black holes with mass M in the range 5×10^{15} g $\lesssim M \lesssim 10^{21}$ g. This is a range still allowed by existing observations.

It is, of course, an open question whether an inflationary model with broken scale invariance is realized in nature. Apart from this, it is also open whether the assumption of Gaussian fluctuations holds; primordial fluctuations may have a non-Gaussian statistics.

The PBH formation discussed here arises from classical stochastic density fluctuations. However, these fluctuations originate from quantum fluctuations and become classical only through decoherence (Section 10.2.3). There is also a small range of wavelengths for which decoherence is inefficient and for which the fluctuations thus stay quantum until they re-enter the Hubble scale. It is an open and intriguing question whether such fluctuations could form a 'quantum black hole' and what the properties of such an object could be.

The question of whether primordial black holes really exist in nature or whether they have only theoretical significance has not yet been settled. Their presence would be of an importance that can hardly be overestimated. They would give a unique opportunity to study the quantum effects of black holes and could yield the crucial key to the construction of a final theory of quantum gravity. The search for them remains exciting, both theoretically and observationally.

Further reading: Barbero and Villaseñor (2010), Bekenstein (2004), Hawking (1975), Kuchař (1994).

8 Quantum cosmology

8.1 Minisuperspace models

8.1.1 General introduction

As we have discussed at length in Chapters 5 and 6, all information about canonical quantum gravity lies in the constraints (apart from possible surface terms). These constraints assume, in the Dirac approach, the form of conditions for physically allowed wave functionals. In quantum geometrodynamics (Chapter 5), the wave functional depends on the three-metric, while in quantum connection or quantum loop dynamics (Chapter 6), it depends on a non-Abelian connection or a Wilson-loop-type variable.

A common feature for all variables is that the quantum constraints are difficult, if not impossible, to solve. In classical GR, the field equations often become tractable if *symmetry reductions* are performed; one can, for example, impose spherical symmetry, axial symmetry, or homogeneity. This often corresponds to interesting physical situations: stationary black holes are spherically or axially symmetric (Section 7.1), while the Universe as a whole can be approximated by homogeneous and isotropic models. The idea is to apply a similar procedure in the quantum theory. One may wish to make a symmetry reduction at the classical level and to quantize only a restricted set of variables. The quantization of black holes discussed in Chapter 7 is a prominent example. The problem is that such a reduction violates the uncertainty principle, since degrees of freedom are neglected together with the corresponding momenta. Nevertheless, the reduction may be an adequate approximation in many circumstances. In quantum mechanics, for example, the model of a 'rigid top' is a good approximation as long as other degrees of freedom remain unexcited due to energy gaps. Such a situation may also hold for quantum gravity; cf. Kuchař and Ryan (1989) for a discussion of this situation in a quantum-cosmological context. In the dynamical-triangulation approach discussed in Section 2.2.6, for example, one can derive an effective cosmological action from the full path integral (Ambjørn *et al.* 2005).

Independently of this question of whether the resulting models are realistic or not, further reasons support the study of dimensionally reduced models. First, they can play the role of toy models to study conceptual issues that are independent of the number of variables. Examples are the problem of time, the role of observers, and the emergence of a classical world; cf. Chapter 10. Second, they can provide a means to study mathematical questions such as the structure of the wave equation and the implementation of boundary conditions. Third, one can compare alternative quantization schemes in the context of simple models. This concerns, in particular, a comparison between reduced and Dirac quantization.

In quantum geometrodynamics, the wave functional Ψ is defined—apart from non-gravitational degrees of freedom—on Riem Σ, the space of all three-metrics. The presence of the diffeomorphism constraints guarantees that the true configuration space is 'superspace', that is, the space Riem Σ/Diff Σ. By restricting the infinitely many degrees of freedom of superspace to only a finite number, one arrives at a finite-dimensional configuration space called *minisuperspace*. If the number of restricted variables is still infinite, the term 'midisuperspace' is used. The spherically symmetric case discussed in the last chapter is an example of midisuperspace. Since the most important example in the case of finitely many degrees of freedom is cosmology, examples of minisuperspace are usually applied to *quantum cosmology*, that is, the application of quantum theory to the Universe as a whole. In fact, the term 'quantum cosmology' is used almost exclusively to describe drastically reduced models. The present chapter deals with quantum cosmology. For simplicity, we shall restrict our attention mostly to quantum geometrodynamics; the application of loop quantum gravity (cf. Chapter 6) to cosmology is called *loop quantum cosmology* and is discussed in Section 8.4 below.

The importance of quantum theory for an understanding of the origin of the universe was emphasized by Georges Lemaître in the context of his *atome primitif*; see Lemaître (1958). As early as 1931, he wrote (Lemaître 1931)

> If the world has begun with a single quantum, the notions of space and time would altogether fail to have any meaning at the beginning ... If this suggestion is correct, the beginning of the world happened a little before the beginning of space and time.

However, he did not consider the quantization of space–time itself. The idea that the Universe as a whole is the result of a 'vacuum fluctuation' in quantum field theory can be traced back at least to Tryon (1973).

Independently of any quantum theory of gravity, there are general consistency arguments that require the application of quantum theory to the Universe as a whole. Namely, macroscopic quantum systems are strongly coupled to their natural environment; cf. the discussion in Chapter 10. Since the environment is again coupled to its environment, and so on, the only strictly closed system in the quantum-theoretical sense is the Universe as a whole. This leads to quantum cosmology, whatever interactions are involved. However, since gravity is the dominant interaction on cosmic scales, a quantum theory of gravity is needed as the formal framework for quantum cosmology.

The first quantum-cosmological model based on quantum gravity was presented, together with its semiclassical approximation, in DeWitt (1967*a*). It dealt with the homogeneous and isotropic case. The extension to anisotropic models (in particular, Bianchi models) was performed by Misner; cf. Misner (1972) and Ryan (1972). Kuchař (1971) made the extension to midisuperspace and discussed the quantization of cylindrical gravitational waves; see also Ashtekar and Pierri (1996). Classically, in the general case of inhomogeneous models, different spatial points seem to decouple near a big-bang singularity for general solutions of the Einstein equations (Belinskii *et al.* 1982). Such solutions consist of a collection of homogeneous spaces described, for example, by a 'mixmaster universe' (in which the universe behaves like a particle in a time-dependent potential wall, with an infinite sequence of bounces). For this reason, the use of minisuperspace models may even provide a realistic description of the universe near its classical singularity.

The usual symmetry reduction proceeds as follows; see, for example, Torre (1999). One starts from a classical field theory and specifies the action of a group with respect to which the fields are supposed to be invariant. A prominent example is the rotation group. One then constructs the invariant ('reduced') fields and evaluates the field equations for them. An important question is whether there is a shortcut in the following way. Instead of reducing the field equations, one might wish to reduce the Lagrangian first and then derive the reduced field equations directly from it. (Alternatively, this can be attempted at the Hamiltonian level.) This would greatly simplify the procedure, but in general it is not possible: reduction of the Lagrangian is equivalent to reduction of the field equations only in special situations. When do such situations occur? In other words, when do critical points of the reduced action define critical points of the full action? Criteria for this *symmetric criticality principle* were developed by Palais (1979). If restriction is made to local Lagrangian theories, one can specify such criteria more explicitly (Torre 1999; Fels and Torre 2002). Instead of spelling out the general conditions, we focus on three cases that are relevant for us:

- The conditions are always satisfied for a compact symmetry group, that is, the important case of *spherical symmetry* obeys the symmetric criticality principle.
- In the case of homogeneous cosmological models, the conditions are satisfied if the structure constants c^c_{ab} of the isometry group satisfy $c^b_{ab} = 0$. Therefore, Bianchi-type-A cosmological models and the Kantowski–Sachs universe can be treated via a reduced Lagrangian. For Bianchi-type-B models, the situation is more subtle (cf. MacCallum (1979) and Ryan and Waller (1997)).
- The symmetric criticality principle also applies to cylindrical or toroidal symmetry reductions (which are characterized by two commuting Killing vector fields). The reduced theories can be identified with parametrized field theories on a flat background. With such a formal identification it is easy to find solutions. Quantization can then be understood as quantization on a fixed background with arbitrary foliation into Cauchy surfaces. In two space–time dimensions, where these reduced models are effectively defined, time evolution is unitarily implementable along arbitrary foliations. This ceases to hold in higher dimensions; cf. Giulini and Kiefer (1995), Helfer (1996), and Torre and Varadarajan (1999).

In the case of homogeneous models, the wave function is of the form $\psi(q^i)$, $i = 1, \ldots, n$; that is, it is of a 'quantum-mechanical' type. In this section, we follow a pragmatic approach and discuss the differential equations for the wave functions, together with appropriate boundary conditions. A general discussion of boundary conditions will be presented in Section 8.3 below. Extensive reviews of quantum cosmology include Halliwell (1991), Wiltshire (1996), and Coule (2005); a short introduction is Berger (1993).

8.1.2 Quantization of a Friedmann universe

As an example, we shall treat in some detail the case of a closed Friedmann universe with a massive scalar field; cf. Kiefer (1988) and Halliwell (1991). Classically, the model is thus characterized by a scale factor $a(t)$ and a homogeneous field $\phi(t)$ with mass m. In the quantum theory, the classical time parameter t is absent, and the system is fully characterized by a wave function $\psi(a,\phi)$. The ansatz for the classical line element is

$$\mathrm{d}s^2 = -N^2(t)\,\mathrm{d}t^2 + a^2(t)\,\mathrm{d}\Omega_3^2, \tag{8.1}$$

where $\mathrm{d}\Omega_3^2 = \mathrm{d}\chi^2 + \sin^2\chi(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)$ is the standard line element on S^3. A special foliation has thus been chosen in order to capture the symmetries of this model. For this reason, no shift vector appears, only the lapse function N. The latter naturally occurs in combination with $\mathrm{d}t$, expressing the classical invariance under reparametrizations of the time parameter; see Chapter 3.

The three-metric h_{ab} is fully specified by the scale factor a. The second fundamental form (4.48) reads

$$K_{ab} = \frac{1}{2N}\frac{\partial h_{ab}}{\partial t} = \frac{\dot{a}}{aN}h_{ab} \tag{8.2}$$

here. Its trace is

$$K := K_{ab}h^{ab} = \frac{3\dot{a}}{Na}, \tag{8.3}$$

which is thus proportional to the Hubble parameter $\dot{a}/a$. Its inverse, K^{-1}, is called the 'extrinsic time' for this reason; cf. Section 5.2.

Since the model fulfils the symmetric criticality principle, we can insert the ansatz (8.1) directly into the Einstein–Hilbert action (1.1) and derive the Euler–Lagrange equations from the reduced action. For the surface term in (1.1), one obtains

$$\frac{1}{8\pi G}\int \mathrm{d}^3x\ \sqrt{h}K = \frac{3}{8\pi G}\int \mathrm{d}^3x\ \sqrt{h}\frac{\dot{a}}{Na}.$$

Inserting

$$\sqrt{h}\ \mathrm{d}^3x = a^3\sin^2\chi\sin\theta\ \mathrm{d}\chi\,\mathrm{d}\theta\,\mathrm{d}\varphi,$$

one finds that this surface term is cancelled by a term that appears after partial integration from the first term in (1.1). This is how the general surface term is constructed. More explicitly, one has the following for the curvature scalar:

$$R = \frac{6}{N^2}\left(-\frac{\dot{N}\dot{a}}{Na} + \frac{\ddot{a}}{a} + \left[\frac{\dot{a}}{a}\right]^2\right) + \frac{6}{a^2}.$$

Partial integration of the second term in the parentheses cancels both the surface term and the first term and changes the sign of the third term.

Integrating over d^3x and choosing units such that $2G/3\pi = 1$, one obtains from (1.1) the gravitational part of the 'minisuperspace action',

$$S_{\mathrm{g}} = \frac{1}{2}\int \mathrm{d}t\ N\left(-\frac{a\dot{a}^2}{N^2} + a - \frac{\Lambda a^3}{3}\right). \tag{8.4}$$

The matter action reads, after a rescaling $\phi \to \phi/\sqrt{2}\pi$ and $m^2/\hbar^2 \to m^2$,

$$S_{\rm m} = \frac{1}{2}\int {\rm d}t\ Na^3\left(\frac{\dot\phi^2}{N^2} - m^2\phi^2\right). \tag{8.5}$$

The full minisuperspace action can then be written in the form

$$S = S_{\rm g} + S_{\rm m} \equiv \int {\rm d}t\ L(q,\dot q) \equiv \int {\rm d}t\ N\left(\frac{1}{2}G_{AB}\frac{\dot q^A \dot q^B}{N^2} - V(q)\right), \tag{8.6}$$

where q is a shorthand for $q^1 := a$ and $q^2 := \phi$. The minisuperspace DeWitt metric (the analogue of (4.23) in minisuperspace) can be read off from (8.4) and (8.5); it reads explicitly

$$G_{AB} = \begin{pmatrix} -a & 0 \\ 0 & a^3 \end{pmatrix},$$

with $\sqrt{-G} = a^2$, where G denotes the determinant of the metric here. The indefinite nature of the kinetic term is evident.

Following the general procedure of Section 4.2, one starts with the definition of the canonical momenta. This includes the primary constraint

$$p_N = \frac{\partial L}{\partial \dot N} \approx 0$$

and the dynamical momenta

$$p_a = \frac{\partial L}{\partial \dot a} = -\frac{a\dot a}{N}\,, \quad p_\phi = \frac{\partial L}{\partial \dot\phi} = \frac{a^3\dot\phi}{N}. \tag{8.7}$$

The Hamiltonian is then given by

$$\begin{aligned} H &= p_N\dot N + p_a\dot a + p_\phi\dot\phi - L \\ &= \frac{N}{2}\left(-\frac{p_a^2}{a} + \frac{p_\phi^2}{a^3} - a + \frac{\Lambda a^3}{3} + m^2a^3\phi^2\right) \\ &\equiv N\left(\tfrac{1}{2}\ G^{AB}p_Ap_B + V(q)\right), \end{aligned} \tag{8.8}$$

where G^{AB} is the inverse DeWitt metric. The explicit form of the potential is

$$V(q) := V(a,\phi) = \frac{1}{2}\left(-a + \frac{\Lambda a^3}{3} + m^2a^3\phi^2\right) =: \frac{1}{2}\left(-a + a^3\mathcal{V}(\phi)\right).$$

For a general (minimally coupled) field, one has to insert the corresponding potential term into $\mathcal{V}(\phi)$. One recognizes that the Λ-term appears on the same footing, so one might also have an effective Λ-term coming from a matter potential. This is the typical case for inflationary scenarios of the early universe.

For convenience, we give here the expressions for p_a and H for a general spatial coordinate volume V_0 and value of G. In the case above, $V_0 = 2\pi^2$ (the volume of S^3), but one can also have a flat model with a finite volume. This is why one may also allow the curvature parameter k for the Friedmann universe (which above and below assumes the value $k = +1$) to assume the value zero. The expressions read

$$p_a = -\frac{3V_0}{4\pi G}\frac{a\dot{a}}{N} \tag{8.9}$$

and

$$H = \frac{N}{2}\left(-\frac{2\pi G}{3}\frac{p_a^2}{V_0 a} + \frac{p_\phi^2}{a^3} - \frac{3}{8\pi G}V_0 k a + \frac{V_0 \Lambda a^3}{8\pi G} + m^2 a^3 \phi^2\right). \tag{8.10}$$

(The field ϕ has been rescaled here according to $\phi \to \phi/\sqrt{V_0}$.) In the following, we shall return to the original choice for G, V_0, and k, but we shall make use of the general expressions in Section 8.4.

As in the general canonical formalism, the preservation of the primary constraint, $\{p_N, H\} \approx 0$, leads to the Hamiltonian constraint $H \approx 0$. Due to the ansatz (8.1), no diffeomorphism constraints appear. At the Lagrangian level, $\{p_N, H\} \approx 0$ corresponds to the Friedmann equation, which for $N = 1$ reads

$$\dot{a}^2 = -1 + a^2\left(\dot{\phi}^2 + \frac{\Lambda}{3} + m^2\phi^2\right). \tag{8.11}$$

Variation of (8.6) with respect to ϕ yields

$$\ddot{\phi} + \frac{3\dot{a}}{a}\dot{\phi} + m^2\phi = 0. \tag{8.12}$$

The classical equations can only be solved analytically for $m = 0$. In that case one has $p_\phi = a^3\dot{\phi} = \text{constant} =: \mathcal{K}$, leading for $\Lambda = 0$ to

$$\phi(a) = \pm\frac{1}{2}\text{arcosh}\frac{\mathcal{K}}{a^2}. \tag{8.13}$$

If $m \neq 0$, a typical solution in configuration space behaves as follows: beginning at a point with small a and $\phi \neq 0$, the trajectory approaches the a-axis and starts to oscillate around it. This model is often used in the context of 'chaotic inflation' (cf. Linde 1990), because the part of the trajectory approaching $\phi = 0$ corresponds to inflationary expansion with respect to the coordinate time t. For a closed Friedmann universe, the trajectory reaches a maximum and recollapses.

Quantization proceeds through implementation of $H \approx 0$ as a condition on the wave function. As in the general theory, there is the freedom to choose the factor ordering. The suggestion put forward here is a choice that leads to a kinetic term that is invariant under transformations in configuration space. In fact, this is already the situation in the ordinary Schrödinger equation. It corresponds to the substitution

$$\begin{aligned} G^{AB}p_A p_B \longrightarrow -\hbar^2\nabla^2_{\text{LB}} &:= -\frac{\hbar^2}{\sqrt{-G}}\partial_A(\sqrt{-G}G^{AB}\partial_B) \\ &= \frac{\hbar^2}{a^2}\frac{\partial}{\partial a}\left(a\frac{\partial}{\partial a}\right) - \frac{\hbar^2}{a^3}\frac{\partial^2}{\partial\phi^2}, \end{aligned} \tag{8.14}$$

where the 'Laplace–Beltrami operator' ∇_{LB} is the covariant generalization of the Laplace operator. With this factor ordering, the Wheeler–DeWitt equation, $\hat{H}\psi(a,\phi) = 0$, reads

$$\frac{1}{2}\left(\frac{\hbar^2}{a^2}\frac{\partial}{\partial a}\left(a\frac{\partial}{\partial a}\right) - \frac{\hbar^2}{a^3}\frac{\partial^2}{\partial\phi^2} - a + \frac{\Lambda a^3}{3} + m^2 a^3\phi^2\right)\psi(a,\phi) = 0. \tag{8.15}$$

This equation assumes a particularly simple form if the variable $\alpha := \ln a$ is used instead of a:[1]

$$\frac{\mathrm{e}^{-3\alpha}}{2}\left(\hbar^2\frac{\partial^2}{\partial\alpha^2} - \hbar^2\frac{\partial^2}{\partial\phi^2} - \mathrm{e}^{4\alpha} + \mathrm{e}^{6\alpha}\left[m^2\phi^2 + \frac{\Lambda}{3}\right]\right)\psi(\alpha,\phi) = 0. \tag{8.16}$$

The variable α ranges from $-\infty$ to ∞, and there are thus no problems in connection with a restricted range of the configuration variable.

Equation (8.16) has the form of a Klein–Gordon equation with a 'time-and space-dependent' mass term, that is, with a non-trivial potential term given by

$$V(\alpha,\phi) = -\mathrm{e}^{4\alpha} + \mathrm{e}^{6\alpha}\left[m^2\phi^2 + \frac{\Lambda}{3}\right]. \tag{8.17}$$

Since this potential can also become negative, the system can develop 'tachyonic' behaviour with respect to the minisuperspace lightcone defined by the kinetic term in (8.16). This does not lead to any inconsistency, since one is dealing with a configuration space here, not with space–time.

The minisuperspace configuration space of this model (the space spanned by α and ϕ) is conformally flat. This is an artefact of two dimensions. In general, the configuration space is curved. Therefore, the presence of an additional factor-ordering term of the form $\eta G\hbar^2\mathcal{R}/a$ in (8.15) is possible, where η is a pure number and $\mathcal{R}$ is the Ricci scalar of the configuration space. It has been argued (Misner 1972, Halliwell 1988, Christodoulakis and Korfiatis 1991) that the choice

$$\eta = \frac{d-2}{8(d-1)},$$

where $d \neq 1$ is the dimension of minisuperspace, is preferred. This choice follows from the demand for the invariance of the Wheeler–DeWitt equation under conformal transformations of the DeWitt metric.

One recognizes from (8.16) that α plays the role of an 'intrinsic time'—the variable that comes with the opposite sign in the kinetic term.[2] Since the potential (8.17) obeys $V(\alpha,\phi) \neq V(-\alpha,\phi)$, there is no invariance with respect to reversal of *intrinsic*

[1]More precisely, one should define $\alpha := \ln a/a_0$ with some reference scale a_0, e.g. $a_0 = l_{\mathrm{P}}$. For simplicity, we set $a_0 = 1$ here.

[2]In two-dimensional configuration spaces, only the relative sign seems to play a role. However, in higher-dimensional minisuperspaces, it becomes clear that the variable connected with the *volume* of the universe is the time-like variable; see Section 5.2.2.

time. This is of crucial importance to understanding the origin of irreversibility; cf. Section 10.2. Moreover, writing (8.16) in the form

$$-\hbar^2 \frac{\partial^2}{\partial \alpha^2}\psi \equiv h_\alpha^2 \psi, \tag{8.18}$$

the 'reduced Hamiltonian' h_α is not self-adjoint, so there is no unitary evolution with respect to α. This, however, is not a problem, since α is not an external time. Unitarity with respect to an intrinsic time is not expected to hold.

Models such as the one above can thus serve to illustrate the difficulties that arise in the approaches of 'reduced quantization' discussed in Section 5.2 (Blyth and Isham 1975). Choosing classically $a = t$ and solving $H \approx 0$ for p_a leads to $p_a + h_a \approx 0$ and, therefore, to a reduced Hamiltonian h_a equivalent to the one in (8.18),

$$h_a = \pm\sqrt{\frac{p_\phi^2}{t^2} - t^2 + \frac{\Lambda t^4}{3} + m^2\phi^2 t^4}. \tag{8.19}$$

One can recognize all the problems that are connected with such a formulation—explicit t-dependence of the Hamiltonian, no self-adjointness, and a complicated expression. One can make alternative choices for t—either $p_a = t$ ('extrinsic time'), $\phi = t$ ('matter time'), or a mixture of all these. This non-uniqueness is an expression of the 'multiple-choice problem' mentioned in Section 5.2. Due to these problems, we shall restrict our attention to the discussion of the Wheeler–DeWitt equation and shall not follow the reduced approach any further.

A simple special case of (8.16) is obtained if we set $\Lambda = 0$ and $m = 0$. Again setting $\hbar = 1$ also, we get

$$\left(\frac{\partial^2}{\partial \alpha^2} - \frac{\partial^2}{\partial \phi^2} - \mathrm{e}^{4\alpha}\right)\psi(\alpha, \phi) = 0. \tag{8.20}$$

A separation ansatz leads immediately to the special solution

$$\psi_k(\alpha, \phi) = \mathrm{e}^{-\mathrm{i}k\phi} K_{\mathrm{i}k/2}\left(\frac{\mathrm{e}^{2\alpha}}{2}\right), \tag{8.21}$$

where $K_{\mathrm{i}k/2}$ denotes a Bessel function. This solution has been chosen in order to fulfil the boundary condition that $\psi \to 0$ as $\alpha \to \infty$. A general solution with this boundary condition can be found by performing a superposition with a suitable amplitude $A(k)$,

$$\psi(\alpha, \phi) = \int_{-\infty}^{\infty} \mathrm{d}k \; A(k)\psi_k(\alpha, \phi). \tag{8.22}$$

Taking, for example, a Gaussian centred at $k = \bar{k}$ with width b,

$$A(k) = \frac{1}{\sqrt{\sqrt{\pi} b}} \exp\left(-\frac{(k - \bar{k})^2}{2b^2}\right), \tag{8.23}$$

we obtain the following wave-packet solution for the real part of the wave function (cf. Kiefer 1988):

$$\begin{aligned}\mathrm{Re}\ \psi(\alpha,\phi) \approx\ & c_{\bar{k}} \cos f_{\bar{k}}(\alpha,\phi) \mathrm{e}^{-(b^2/2)\left(\phi+1/2\ \mathrm{arcosh}(\bar{k}/\mathrm{e}^{2\alpha})\right)^2} \\ & + c_{\bar{k}} \cos g_{\bar{k}}(\alpha,\phi) \mathrm{e}^{-(b^2/2)\left(\phi-1/2\ \mathrm{arcosh}(\bar{k}/\mathrm{e}^{2\alpha})\right)^2} . \end{aligned} \tag{8.24}$$

The functions f and g are given explicitly in Kiefer (1988) and are not needed here. The wave function (8.24) is a sum of two (modulated) Gaussians of width b^{-1}; they are symmetric with respect to $\phi = 0$ and follow the classical path given by (8.13) with $\mathcal{K} = \bar{k}$.

This example demonstrates an important feature of the quantum theory (see Fig. 8.1). In the classical theory, the 'recollapsing' part of the trajectory in configuration space can be considered as the deterministic successor of the 'expanding' part. However, in the quantum theory, the 'returning' part of the wave packet must be present 'initially' with respect to the intrinsic time (the scale factor a of the universe) in order to yield a wave tube that follows the classical trajectory. This ensures destructive interference near the classical turning point in order to avoid exponentially growing components of the wave function at large a.

Whereas in the above example it is possible to construct wave packets that follow the classical trajectory without dispersion, this is not possible in the general case. For $m \neq 0$, one finds that the demand for $\psi \to 0$ as $a \to \infty$ is in conflict with the existence of narrow wave packets all along the classical trajectory (Kiefer 1988). Consider the Wheeler–DeWitt equation (8.16) for $\Lambda = 0$,

$$\left(\frac{\partial^2}{\partial \alpha^2} - \frac{\partial^2}{\partial \phi^2} - \mathrm{e}^{4\alpha} + \mathrm{e}^{6\alpha} m^2 \phi^2 \right) \psi(\alpha,\phi) = 0. \tag{8.25}$$

Within a Born–Oppenheimer type of approximation, one can make the following ansatz for the wave packet:

$$\psi(a,\phi) = \sum_n c_n(\alpha)\, \varphi_n \left(\sqrt{m\mathrm{e}^{3\alpha}}\phi \right) , \tag{8.26}$$

where the φ_n denote the usual eigenfunctions of the harmonic oscillator. From (8.25), one gets the following effective potentials for the $c_n(\alpha)$:

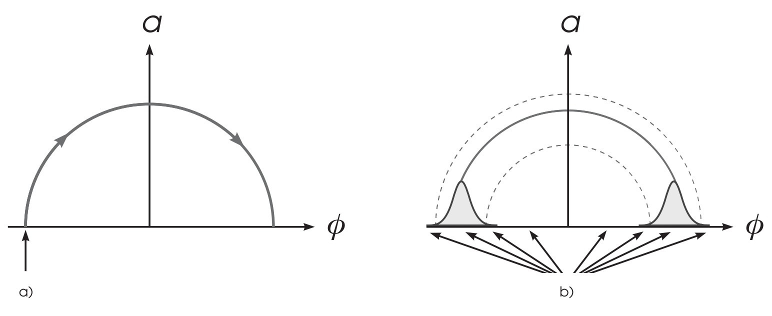

Fig. 8.1 (a) *Classical theory.* One can give, for example, initial conditions as indicated by the left arrow. The recollapsing part of the trajectory is then the deterministic successor of the expanding part. (b) *Quantum theory.* Initial conditions are given on $a =$ constant. The 'recollapsing' part of the wave packet must be present 'initially'.

$$V_n(\alpha) = \tfrac{1}{2}\left(-\mathrm{e}^{4\alpha} + (2n+1)m\mathrm{e}^{3\alpha}\right). \tag{8.27}$$

These potentials become negative for large enough α. In the classical theory, this means that trajectories are drawn into the region with negative V_n and are reflected. In the quantum theory, it means that the wave function is a superposition of exponentially increasing and decreasing solutions. In order to fulfil the 'final condition' $\psi \to 0$ as $a \to \infty$, the exponentially *decreasing* solution must be chosen. The n-dependent reflection expressed by (8.27) leads to an unavoidable *spreading* of the wave packet (Kiefer 1988). This means that the semiclassical approximation does not hold all along the expanding and recollapsing parts of the classical trajectory. How, then, does classical behaviour emerge? The answer is provided by adding other degrees of freedom (Section 8.2). These can act as a kind of environment for the minisuperspace variables a and ϕ and force them to behave classically. This process of *decoherence* is discussed in Section 10.1. It must also be emphasized that wave packets are always understood here as corresponding to *branches* (components) of the full wave function (representing quasi-classical histories), and not to the full wave function itself.

Another example is a closed Friedmann universe with a *non-minimally* coupled scalar field. In the general case, the Wheeler–DeWitt equation can become elliptic instead of hyperbolic for a certain range of field values; see Kiefer (1989). This would modify the 'initial-value problem' in quantum cosmology. No elliptic region occurs, for example, in the simplest case of a conformally coupled field ϕ. Choosing units such that $8\pi G = 1$ and performing a field redefinition

$$\phi \to \chi = \frac{\sqrt{2}\pi a\phi}{6},$$

one obtains the following Wheeler–DeWitt equation (Zeh 1988, Kiefer 1990):

$$\hat{H}\psi(a,\chi) \equiv (-H_a + H_\chi)\psi \equiv \left(\frac{\partial^2}{\partial a^2} - \frac{\partial^2}{\partial \chi^2} - a^2 + \chi^2\right)\psi = 0. \tag{8.28}$$

This has the form of an 'indefinite oscillator'—two harmonic oscillators distinguished by a relative sign in the Hamiltonian. The model defined by (8.28) is the simplest non-trivial model in quantum cosmology. (It also arises from scalar–tensor theories of gravity; cf. Lidsey (1995).) Its classical solutions are represented by Lissajous ellipses confined to a rectangle in configuration space. The corresponding wave packets are obtained if a normalization condition is imposed with respect to a and χ. In this way, one obtains an ordinary quantum-mechanical Hilbert space. The wave packets follow the classical Lissajous ellipses without dispersion here (Kiefer 1990; Gousheh and Sepangi 2000). From (8.28), one finds

$$H_a\varphi_n(a)\varphi_n(\chi) = H_\chi\varphi_n(a)\varphi_n(\chi),$$

where the φ_n again denote the usual harmonic-oscillator eigenfunctions. A wave-packet solution can then be constructed according to

$$\psi(a,\chi) = \sum_n A_n\varphi_n(a)\varphi_n(\chi) = \sum_n A_n \frac{H_n(a)H_n(\chi)}{2^n n!}\mathrm{e}^{-a^2/2-\chi^2/2}, \tag{8.29}$$

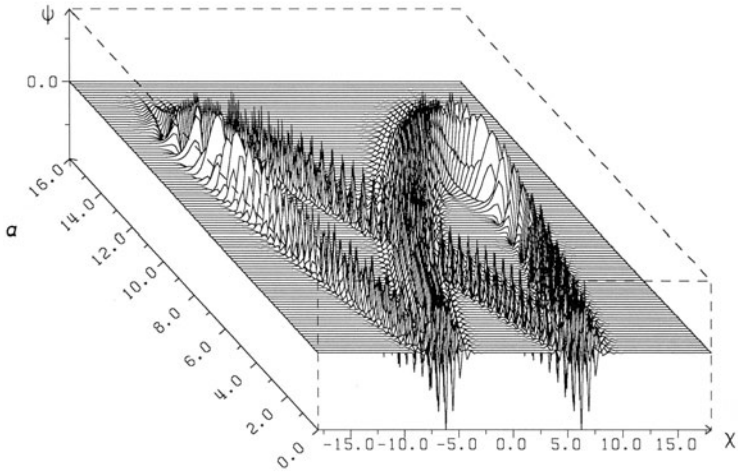

Fig. 8.2 Wave packet corresponding to a classical Lissajous figure (only the region $a \geq 0$ is shown). From Kiefer (1990). © 1990 by Elsevier B.V.

with suitable coefficients A_n. From the properties of the Hermite polynomials, it is evident that the wave packet must satisfy the 'initial condition' $\psi(0, \chi) = \psi(0, -\chi)$. The requirement of normalizability thus restricts the possible initial conditions. A particular example of such a wave-packet solution is depicted in Fig. 8.2. One recognizes a superposition of (half of) two Lissajous ellipses.

In a more general oscillator model one would expect a and χ to have different frequencies, that is, instead of the potential $-a^2 + \chi^2$, one would have $-\omega_a^2 a^2 + \omega_\chi^2 \chi^2$. The demand for normalizability would then entail the commensurability condition

$$\frac{\omega_\chi}{\omega_a} = \frac{2n_a + 1}{2n_\chi + 1}, \tag{8.30}$$

where n_a and n_χ are integer numbers. Thus, one gets from normalizability a restriction on the 'coupling constants' of this model. It is imaginable that such conditions may also hold in the full theory, for example for the cosmological constant.[3]

As Page (1991) has demonstrated, one can map various minisuperspace models into each other. He has also presented numerous classical and quantum solutions for these models. For example, one can find a map between the cases of minimally and non-minimally coupled scalar fields. The fields have to be rescaled and, what is most important, they differ in the range of allowed values. As mentioned above, this has consequences for the initial-value problem in quantum gravity (Kiefer 1989). A particular example is the map between the models defined by (8.28) and (8.15) for $m = 0$ and $\Lambda = 0$, respectively. The respective wave functions, however, have different domains

[3]Quantization conditions for the cosmological constant have been derived from a particular minisuperspace model in a rigorous way by Gerhardt (2009). Such quantization conditions may also arise in string theory; cf. Bousso and Polchinski (2000) and Feng *et al.* (2001).

for their argument. Another map relates (8.28) to a massive Klein–Gordon equation in the (1+1)-dimensional 'Rindler wedge'.

Up to now, we have addressed quantum cosmology in the framework of geometrodynamic variables with (mostly) minimally coupled fields and without supersymmetry (SUSY). It is possible to extend the above discussion to the presence of SUSY, to effective actions coming from string theory ('quantum string cosmology'), and to quantum connection and loop dynamics. In the following, we shall add some remarks on each of these frameworks.

Canonical supergravity (SUGRA) was discussed in Section 5.3.6. Again, one can restrict the corresponding action to spatially homogeneous models, in the simplest case to a Friedmann universe. After this specialization, only the spinor indices remain. In this way, one can say that full SUGRA with $N = 1$ leads to an effective minisuperspace model with $N = 4$ SUSY—one has only four quantum-mechanical generators of SUSY, S_A $(A = 1, 2)$ and $\bar{S}_{A'}$ $(A' = 1', 2')$; cf. (5.99). As we discussed in Section 5.3.6, it is sufficient to solve the Lorentz and SUSY constraints; the Hamiltonian and diffeomorphism constraints follow through the Dirac-bracket relations. This also holds, of course, in minisuperspace. The Lorentz constraints can be implemented by performing the following decomposition of the wave function Ψ into the fermionic variables, allowing the coefficients to be functions of the bosonic variables:

$$\begin{aligned}\Psi = \mathcal{A}(a,\phi,\bar{\phi}) + \mathcal{B}(a,\phi,\bar{\phi})\psi^A\psi_A + \mathrm{i}\mathcal{C}(a,\phi,\bar{\phi})\chi^A\psi_A \\ +\mathcal{D}(a,\phi,\bar{\phi})\chi^A\chi_A + \mathcal{E}(a,\phi,\bar{\phi})\psi^A\psi_A\chi^A\chi_A.\end{aligned} \tag{8.31}$$

Here, a is the scale factor of the Friedmann universe, ϕ is a complex scalar field, χ_A is its fermionic superpartner, and ψ_A is the homogeneous degree of freedom coming from the gravitino field. The configuration space of Ψ is thus given by $(\chi_A, \psi_A, a, \phi, \bar{\phi})$. The SUSY constraints then yield a system of coupled first-order differential equations for the coefficient functions $\mathcal{A}$–$\mathcal{E}$. They may be solved in various situations, and particular boundary conditions (Section 8.3) may be imposed. For a detailed discussion, we refer to Moniz (2003, 2010) and the references therein.

The method of canonical quantization can also be applied to actions different from the Einstein–Hilbert action for gravity. Popular examples are given by effective actions from string theory; cf. Section 9.2. They contain, besides the metric, additional fields in the gravitational sector, notably the dilaton field and axion fields. They can act as the starting point for both classical and quantum cosmology; see Gasperini and Veneziano (2003) for an extensive review. String theory possesses an invariance with respect to 'duality transformations' (Chapter 9). This is reflected in the context of Friedmann cosmology by the duality transformation $a \to a^{-1}$ and $t \to -t$. Thereby, one can relate different solutions to each other. Some classical solutions of particular interest are the 'pre-big-bang solutions' for $t < 0$ that are duality-related to 'post-big-bang solutions' for $t > 0$. The former can describe an accelerated (superinflationary) phase driven by the kinetic energy of the dilaton. The latter can describe an expanding decelerated phase, which is interpreted as the standard evolution of the radiation-dominated universe. The hope is that one can connect the two phases and thereby avoid the usual fine-tuning problems of inflationary cosmology. In addition, one can hope to generate primordial gravitational waves with higher amplitude, which could possibly be detected by the space interferometer LISA. However, the transition between the pre- and post-big-bang solutions proceeds through a regime of a classical singularity (strong coupling and high curvature). This is known as the 'graceful exit problem' (exit from inflation into radiation dominance), which is still an open issue.

The quantum version of string cosmology was first investigated by Bento and Bertolami (1995). The role of boundary conditions with respect to the graceful exit problem was discussed by Dąbrowski and Kiefer (1997). Starting from the tree-level effective action with only the metric and the dilaton, one first makes the following redefinition in minisuperspace:

$$\begin{aligned} \beta &:= \sqrt{3}\ln a\ , \\ \bar{\phi} &:= \phi - 3\ln a - \ln\int \frac{\mathrm{d}^3x}{\lambda_{\mathrm{s}}^3}, \end{aligned}$$

where ϕ is the original dilaton field and $\lambda_{\mathrm{s}} := \sqrt{\alpha'}$ is the string length (cf. Section 3.2). One then finds the following Wheeler–DeWitt equation:

$$\left(-\frac{\partial^2}{\partial\bar{\phi}^2} + \frac{\partial^2}{\partial\beta^2} - \lambda_{\mathrm{s}}^2 V(\beta,\bar{\phi})\mathrm{e}^{-2\bar{\phi}}\right)\psi(\beta,\bar{\phi}) = 0, \tag{8.32}$$

where $V(\beta,\bar{\phi})$ denotes the dilaton potential here. Since no external time t exists in quantum cosmology, it does not make sense to talk of a transition between the pre- and post-big-bang regimes. Boundary conditions must be imposed intrinsically, that is, with respect to the configuration-space variables β and $\bar{\phi}$. By constructing wave-packet solutions to (8.32), one recognizes that the pre- and post-big-bang branches just correspond to different solutions (Dąbrowski and Kiefer 1997). An intrinsic distinction between 'expanding' and 'contracting' solutions is not possible, since the reference phase $\mathrm{e}^{-\mathrm{i}\omega t}$ is lacking (Zeh 1988). This can only be achieved if additional degrees of freedom are introduced and a boundary condition of low entropy is imposed; see Section 10.2.

Quantum-cosmological models can be, and have been, discussed in more general situations. Cavaglià and Moniz (2001), for example, have investigated the Wheeler–DeWitt equation for effective actions inspired by 'M-theory' (Section 9.1). Lidsey (1995) has discussed general scalar–tensor theories in which it turns out that the duality symmetry of the classical action corresponds to hidden $N = 2$ SUSY. Quantum cosmology for relativity in more than three space dimensions also possesses some interesting features (see e.g. Zhuk 1992).

Instead of modifying the gravitational sector, one can also introduce exotic matter degrees of freedom. Phantom fields are one example. They have been invoked as a possible explanation for the observed dark energy in the Universe, and are characterized by a negative kinetic term, for which, in the classical theory, such phantom models exhibit new types of singularities. Among them is the *big rip*, where the matter density diverges at finite times for large a instead of small a. The corresponding quantum scenario is discussed in Dąbrowski *et al.* (2006). It turns out that wave-packet solutions of the Wheeler–DeWitt equation disperse near the region that corresponds to the big-rip singularity. One thus arrives at a genuine quantum region at large scales.

The relevance of such models to the quantum avoidance of the classical singularities is discussed in Section 8.5.

Quantum cosmology can also be discussed using methods of connection or loop dynamics (Chapter 6). Paternoga and Graham (1998), for example, investigated Bianchi IX models with $\Lambda \neq 0$ in the connection representation. They started from the Chern–Simons state (6.8), which is a solution of the Euclidean quantum constraints for the Barbero–Immirzi parameter $\beta = 1$. Through a generalized Fourier transform leading to the metric representation, they were able—using inequivalent contours in the transformation formula—to find various solutions to the Wheeler–DeWitt equation. One can also address quantum cosmology directly in the loop representation. The ensuing scenario of loop quantum cosmology is discussed in Section 8.4.

8.1.3 (2+1)-dimensional quantum gravity

General relativity in 2+1 dimensions[4] is 'trivial' in the sense that there are no local dynamical degrees of freedom. The Riemann tensor depends linearly on the Ricci tensor and thus the vacuum solutions of Einstein's equations either are flat (for $\Lambda = 0$) or have constant curvature (for $\Lambda \neq 0$). A typical feature is the appearance of *conical* structures, which give rise to a non-trivial global geometry. There may be a finite number of degrees of freedom connected with the topology of space. The theory is therefore of a quantum-mechanical nature, and it is for this reason that we have included 2+1 quantum gravity in this chapter on quantum cosmology, although this framework is not necessarily restricted to cosmological applications. General references on (2+1)-dimensional gravity include Carlip (1998), Brown (1988), and Matschull (1995); see also Witten (2007).

The Planck mass in 2+1 dimensions is given by

$$m_{\mathrm{P}}^{(3)} = \frac{c^2}{G} \tag{8.33}$$

and is therefore independent of $\hbar$. Classical GR in 2+1 dimensions thus contains a distinguished mass scale. However, Planck's constant enters the Planck length,

$$l_{\mathrm{P}}^{(3)} = \frac{\hbar G}{c^3}. \tag{8.34}$$

The classical canonical formalism here employs a foliation of three-dimensional space–time into two-dimensional spaces Σ. An important theorem states that any metric on a compact two-space Σ is conformal to a metric of constant curvature.[5] The curvature is positive for the two-sphere S^2 (which has $\mathfrak{g} = 0$), zero for the two-torus T^2 ($\mathfrak{g} = 1$), and negative for $\mathfrak{g} > 1$. The two-metric can thus be written as

$$h_{ab}(\mathbf{x}) = \mathrm{e}^{2\xi(\mathbf{x})}\tilde{h}_{ab}(\mathbf{x}), \tag{8.35}$$

with $\tilde{h}_{ab}(\mathbf{x})$ denoting a metric of constant curvature. The role of the configuration space is played here by the *moduli space* of Σ—for $\mathfrak{g} \neq 1$, the space of metrics with

[4]This is a three-dimensional space with signature $(-,+,+)$.

[5]For open Σ, one has to impose appropriate boundary conditions.

constant curvature modulo diffeomorphisms, and for $\mathfrak{g} = 1$ the space of flat metrics of constant prescribed volume modulo diffeomorphisms. The moduli space has a finite dimension: zero for S^2, 2 for T^2, and $6\mathfrak{g} - 6$ for $\mathfrak{g} > 1$. The theory, therefore, describes a finite-dimensional system described by the moduli parameters. It is interesting that a reduced phase space description is possible if 'York's time' (cf. (5.16)) is used (see Carlip 1998). A Schrödinger equation can be formulated in the reduced space; it is similar to the equation occurring in spherically symmetric systems (Section 7.2).

What can be said about the Wheeler–DeWitt quantization (Carlip 1998)? The situation is simplest in a first-order formulation in which the connection ω_i^a and the zweibein e_i^a are treated as independent variables. From the Wheeler–DeWitt equation, one finds that the wave functional must be a functional of flat connections. This is related to the fact that (2+1)-dimensional GR can be formulated as a Chern–Simons theory for a vector potential with gauge group $\mathrm{ISO}(2,1)$ (Achúcarro and Townsend 1986, Witten 1988). In fact, the first-order action of (2+1)-dimensional GR *is* the Chern–Simons action (6.9).

The second-order formalism is much more complicated. There appear functional derivatives with respect to the scale factor ξ and non-local terms from the solution of the diffeomorphism constraints. Most likely, this approach is inequivalent to reduced quantization, and the Wheeler–DeWitt equation cannot be solved, except perhaps in perturbation theory.

Quantum gravity in 2+1 dimensions provides an example of a theory of the first kind in the sense of Section 1.3: it is a consistent quantum theory of the gravitational field itself. This holds irrespective of the fact that it is non-renormalizable by formal power-counting arguments.

Of special interest both classically and quantum mechanically is the existence of a black-hole solution for $\Lambda < 0$, that is, for asymptotic anti-de Sitter space. The solution is called the 'BTZ black hole' after the work by Bañados *et al.* (1992) and is characterized by mass and angular momentum. The BTZ hole provides a toy model in which one can study the issues of Hawking radiation and entropy; cf. Chapter 7. One can, in particular, give a microscopic interpretation of black-hole entropy by counting the degrees of freedom on the horizon in a Chern–Simons approach (Carlip 1998). A detailed view of the BTZ black hole from the perspective of three-dimensional quantum gravity can be found in Witten (2007).

Interesting features appear if point-like particles (whose existence is allowed in 2+1 gravity) are coupled to gravity; cf. Louko and Matschull (2001), who consider the presence of two such particles. For one massive particle, space–time is the product of a conical space with $\mathbb{R}$, the deficit angle of the cone being given by $8\pi Gm$, where m is the mass. For two particles, one has essentially twice this product, but care has to be taken to implement the condition of asymptotic flatness. The particles do not interact directly (there is no Newtonian force in 2+1 gravity), but indirectly through the cone-like structure of space–time. From Wheeler–DeWitt quantization, one finds features that are expected from a quantum theory of gravity: the two particles cannot get closer together than a certain minimal distance given by a multiple of the Planck length. This corresponds to a discrete structure of space–time. Even if the particles are far apart, it is impossible to localize any single particle below a certain length.

All these effects vanish both for $\hbar \to 0$ and for $G \to 0$. A generalization to many particles has been performed by Matschull (2001). He finds, in particular, hints of a non-commutative structure of space–time.

8.2 Introduction of inhomogeneities

The minisuperspace models discussed in the last section are easy to deal with, but are not sufficient for a realistic description of the universe. This can only be achieved if inhomogeneous degrees of freedom are introduced. Otherwise, one would not have the chance of understanding the emergence of structure in quantum cosmology. In the following, we consider a multipole expansion of the three-metric and a scalar field. We take the universe to be closed, so the expansion is with respect to spherical harmonics on the three-sphere S^3. In order to render the formal treatment manageable, the expansion for the 'higher multipoles' is only performed up to quadratic order in the action. They are thus considered to be small perturbations of the homogeneous background described by a and the homogeneous field ϕ. Since one knows from measurements of the microwave background radiation that the fluctuations were small in the early universe, this approximation may be appropriate for that phase. The multipole expansion is also needed for the description of decoherence (Section 10.1.2). Cosmological perturbations were first studied by Lifshits (1946).

Following Halliwell and Hawking (1985), we make the following ansatz for the three-metric:

$$h_{ab} = a^2(\Omega_{ab} + \epsilon_{ab}), \tag{8.36}$$

where Ω_{ab} denotes the metric on S^3, and the 'perturbation' $\epsilon_{ab}(\mathbf{x}, t)$ is expanded into spherical harmonics,

$$\epsilon_{ab}(\mathbf{x}, t) = \sum_{\{n\}} \left(\sqrt{\frac{2}{3}} a_n(t) \Omega_{ab} Q^n + \sqrt{6} b_n(t) P^n_{ab} + \sqrt{2} c_n(t) S^n_{ab} + 2 d_n(t) G^n_{ab} \right). \tag{8.37}$$

Here, $\{n\}$ stands for the three quantum numbers $\{n, l, m\}$, where $n = 1, 2, 3, \ldots$, $l = 0, \ldots, n-1$, and $m = -l, \ldots, l$. The scalar field is expanded as

$$\begin{aligned} \Phi(\mathbf{x}, t) &= \frac{1}{\sqrt{2}\pi} \phi(t) + \epsilon(\mathbf{x}, t), \\ \epsilon(\mathbf{x}, t) &= \sum_{\{n\}} f_n(t) Q^n. \end{aligned} \tag{8.38}$$

The scalar harmonic functions $Q^n := Q^n_{lm}$ on S^3 are the eigenfunctions of the Laplace operator on S^3,

$$Q^{n\ \ |k}_{lm|k} = -(n^2 - 1) Q^n_{lm}, \tag{8.39}$$

where $|k$ denotes the covariant derivative with respect to Ω_{ab}. The harmonics can be expressed as

$$Q^n_{lm}(\chi, \theta, \phi) = \Pi^n_l(\chi) Y_{lm}(\theta, \phi), \tag{8.40}$$

where the $\Pi^n_l(\chi)$ are the 'Fock harmonics', and $Y_{lm}(\theta, \phi)$ are the standard spherical harmonics on S^2. They are orthonormalized according to

$$\int_{S^3} \mathrm{d}\mu \; Q^n_{lm} Q^{n'}_{l'm'} = \delta^{nn'}\delta_{ll'}\delta_{mm'}, \tag{8.41}$$

where $\mathrm{d}\mu = \sin^2\chi \sin\theta \, \mathrm{d}\chi \, \mathrm{d}\theta \, \mathrm{d}\varphi$. The scalar harmonics are thus a complete orthonormal basis with respect to which each scalar field on S^3 can be expanded. The remaining harmonics appearing in (8.37) are called tensorial harmonics of scalar type (P^n_{ab}), vector type (S^n_{ab}), and tensor type (G^n_{ab}); see Halliwell and Hawking (1985) and the references therein. At the present order of approximation for the higher multipoles (up to quadratic order in the action), the vector harmonics are pure gauge. The tensor harmonics G^n_{ab} describe gravitational waves and are gauge-independent. It is possible to work exclusively with gauge-independent variables (Bardeen 1980), but this is not needed for the following discussion.

Introducing the shorthand notation $\{x_n\}$ for the collection of multipoles a_n, b_b, c_n, d_n, the wave function is defined on an infinite-dimensional configuration space spanned by a (or α) and ϕ (the 'minisuperspace background') and the variables $\{x_n\}$. The Wheeler–DeWitt equation can be decomposed into two parts that refer to first and second derivatives, respectively, in the $\{x_n\}$. The part containing the second derivatives reads (Halliwell and Hawking 1985)

$$\left(H_0 + 2\mathrm{e}^{3\alpha} \sum_n H_n(a, \phi, x_n) \right) \Psi(\alpha, \phi, \{x_n\}) = 0, \tag{8.42}$$

where H_0 denotes the minisuperspace part,

$$H_0 := \left(\frac{\partial^2}{\partial\alpha^2} - \frac{\partial^2}{\partial\phi^2} + \mathrm{e}^{6\alpha} m^2 \phi^2 - \mathrm{e}^{4\alpha} \right) =: \left(\frac{\partial^2}{\partial\alpha^2} - \frac{\partial^2}{\partial\phi^2} + V(\alpha, \phi) \right), \tag{8.43}$$

and H_n is a sum of Hamiltonians referring to the scalar, vector, and tensor parts, respectively, of the modes,

$$H_n = H_n^{(S)} + H_n^{(V)} + H_n^{(T)}.$$

We now make the ansatz

$$\Psi(\alpha, \phi, \{x_n\}) = \psi_0(\alpha, \phi) \prod_{n>0} \psi_n(\alpha, \phi; x_n) \tag{8.44}$$

and insert this into (8.42). Following Kiefer (1987), we get

$$\begin{aligned} &-\frac{\nabla^2\psi_0}{\psi_0} - 2\frac{\nabla\psi_0}{\psi_0} \sum_n \frac{\nabla\psi_n}{\psi_n} - \sum_n \frac{\nabla^2\psi_n}{\psi_n} \\ &- \sum_{n\neq m} \frac{\nabla\psi_n \nabla\psi_m}{\psi_n\psi_m} + V(\alpha, \phi) + 2\mathrm{e}^{3\alpha} \sum_n \frac{H_n\Psi}{\Psi} = 0, \end{aligned}$$

where

$$\nabla := \left(\frac{\partial}{\partial\alpha}, \frac{\partial}{\partial\phi} \right)$$

denotes the 'minisuperspace gradient'. One then gets by separation of variables the two equations

$$-\frac{\nabla^2\psi_0}{\psi_0} + V(\alpha,\phi) = -2f(\alpha,\phi), \tag{8.45}$$

$$-2\frac{\nabla\psi_0}{\psi_0}\sum_n \frac{\nabla\psi_n}{\psi_n} - \sum_n \frac{\nabla^2\psi_n}{\psi_n} - \sum_{n\neq m} \frac{\nabla\psi_n\nabla\psi_m}{\psi_n\psi_m}$$
$$+2\mathrm{e}^{3\alpha}\sum_n \frac{H_n\Psi}{\Psi} = 2f(\alpha,\phi), \tag{8.46}$$

where $f(\alpha,\phi)$ is an arbitrary function. The second equation can be separated further if one imposes additional assumptions. First, we assume that the dependence of the ψ_n on the minisuperspace variables is weak in the sense that

$$\left|\frac{\nabla\psi_0}{\psi_0}\right| \gg \left|\frac{\nabla\psi_n}{\psi_n}\right|, \; n \gg 1.$$

Second, we assume that the terms $\nabla\psi_n/\psi_n$ add incoherently (the 'random-phase approximation'), so that the third term on the left-hand side of (8.46) can be neglected compared with the first term. One then gets from (8.46)

$$-\frac{\nabla\psi_0}{\psi_0}\nabla\psi_n - \frac{1}{2}\nabla^2\psi_n + \mathrm{e}^{3\alpha}\left(\frac{H_n\Psi}{\Psi}\right)\psi_n = \varphi_n(\alpha,\phi)\psi_n,$$

where

$$\sum_n \varphi_n(\alpha,\phi) = f(\alpha,\phi).$$

Since the ψ_n are assumed to vary much less with α and ϕ than ψ_0 does, one would expect the term $\nabla^2\psi_n$ to be negligible. Finally, assuming that in H_n the x_n-derivatives dominate over the ∇-derivatives, one can substitute

$$\frac{H_n\Psi}{\Psi}\psi_n \approx H_n\psi_n.$$

One then arrives at

$$-\frac{\nabla\psi_0}{\psi_0}\nabla\psi_n + \mathrm{e}^{3\alpha}H_n\psi_n = \varphi_n\psi_n. \tag{8.47}$$

The choice $f = 0$ in (8.46) would entail that ψ_0 is a solution of the minisuperspace Wheeler–DeWitt equation. If one chooses in addition $\varphi_n = 0$, (8.47) becomes

$$\mathrm{e}^{-3\alpha}\frac{\nabla\psi_0}{\psi_0}\nabla\psi_n = H_n\psi_n. \tag{8.48}$$

If ψ_0 were of WKB form, $\psi_0 \approx C\exp(\mathrm{i}S_0)$ (with a slowly varying prefactor C), one would get

$$\mathrm{i}\frac{\partial\psi_n}{\partial t} = H_n\psi_n, \tag{8.49}$$

with

$$\frac{\partial}{\partial t} := \mathrm{e}^{-3\alpha} \nabla S_0 \cdot \nabla. \tag{8.50}$$

Equation (8.49) is a Schrödinger equation for the multipoles, its time parameter t being defined by the minisuperspace variables α and ϕ. This 'WKB time' controls the dynamics in this approximation. The above derivation reflects the recovery of the Schrödinger equation from the Wheeler–DeWitt equation as discussed in Section 5.4.

One could also choose the φ_n in such a way as to minimize the variation of the ψ_n along the WKB trajectory,

$$\varphi_n = \mathrm{e}^{3\alpha} \langle H_n \rangle, \tag{8.51}$$

where

$$\langle H_n \rangle := \frac{\int \mathrm{d}x_n\ \psi_n^* H_n \psi_n}{\int \mathrm{d}x_n\ \psi_n^* \psi_n}.$$

Instead of (8.49), one then obtains

$$\mathrm{i}\frac{\partial \psi_n}{\partial t} = \left(H_n - \langle H_n \rangle\right) \psi_n. \tag{8.52}$$

The expectation value can be absorbed into the ψ_n if they are redefined by an appropriate phase factor. The minisuperspace equation (8.45) then reads

$$-\frac{\nabla^2 \psi_0}{\psi_0} + V(\alpha, \phi) = -2\mathrm{e}^{3\alpha} \sum_n \langle H_n \rangle. \tag{8.53}$$

The term on the right-hand side corresponds to the back reaction discussed in Section 5.4. From the point of view of the full wave function Ψ, it is just a matter of the splitting between ψ_0 and the ψ_n.

Let us consider as a particular example the case of the tensor multipoles $\{d_n\}$, which describe gravitational waves. After an appropriate redefinition, the wave functions $\psi_n(\alpha, \phi, d_n)$ obey the Schrödinger equations (Halliwell and Hawking 1985)

$$\mathrm{i}\frac{\partial \psi_n}{\partial t} = \frac{1}{2}\mathrm{e}^{-3\alpha} \left(-\frac{\partial^2}{\partial d_n^2} + (n^2 - 1)\mathrm{e}^{4\alpha} d_n^2 \right) \psi_n. \tag{8.54}$$

This has the form of a Schrödinger equation with a 'time-dependent' frequency given by

$$\nu := \frac{\sqrt{n^2 - 1}}{\mathrm{e}^{\alpha}} \overset{n \gg 1}{\approx} n\mathrm{e}^{-\alpha}. \tag{8.55}$$

The (adiabatic) ground-state solution of this equation is

$$\psi_n \propto \exp\left(-\frac{n^3 d_n^2}{2\nu^2} \right) \exp\left(-\frac{\mathrm{i}}{2} \int^t \mathrm{d}s\ \nu(s) \right)\ ; \tag{8.56}$$

see Halliwell and Hawking (1985) and Kiefer (1987). It plays a role in the discussion of primordial fluctuations in inflationary cosmology.

8.3 Boundary conditions

In this section, we shall address an important question that has been neglected so far: what are the appropriate boundary conditions for the Wheeler–DeWitt equation in quantum cosmology?

Since Newton, it has become customary to separate the description of nature into dynamical laws and initial conditions. The latter are usually considered as artificial and can be fixed by the experimentalist, at least in principle. The situation in cosmology is different. The Universe is unique and its boundary conditions are certainly not at our disposal. It has, therefore, been argued that boundary conditions play a key role in the more fundamental framework of quantum cosmology. It has even been claimed that quantum cosmology *is* the theory of initial conditions (see e.g. Hartle 1997 and Barvinsky 2001). In the following, we shall briefly review various proposals for boundary conditions.

8.3.1 DeWitt's boundary condition

DeWitt (1967*a*) suggested imposition of the boundary condition

$$\Psi\left[{}^{(3)}\mathcal{G}\right] = 0 \tag{8.57}$$

for all three-geometries ${}^{(3)}\mathcal{G}$ associated with 'barriers', for example, singular three-geometries. This could automatically alleviate or avoid the singularities of the classical theory. Ideally, one would hope that a unique solution to the Wheeler–DeWitt equation is obtained after this boundary condition is imposed. Whether this is true remains unsettled. In a sense, the demand for the wave function to go to zero at large scale factors—as has been discussed in Section 8.1 in connection with wave packets—can be interpreted as an implementation of (8.57) in minisuperspace. In the example of the collapsing dust shell in Section 7.4, the wave function tends to zero as $r \to 0$, that is, in the region of the classical singularity. In that case, however, this is not the consequence of a boundary condition but of the dynamics—it is a consequence of unitary time evolution.

The danger with imposing the DeWitt condition (8.57) is that only the trivial function $\Psi \equiv 0$ may survive as a solution to the Wheeler–DeWitt equation $H\Psi = 0$.

8.3.2 No-boundary condition

This proposal goes back to Hawking (1982) and Hartle and Hawking (1983). It is therefore also called the 'Hartle–Hawking proposal'. A central role in its formulation is played by Euclidean path integrals. In fact, the description of black-hole thermodynamics by such path integrals was one of Hawking's original motivations for introducing the proposal; cf. Hawking (1979). The 'no-boundary condition' states that for a *compact* three-dimensional space Σ, the wave function Ψ is given by the sum over all compact Euclidean four-geometries of all topologies that have Σ as their *only* boundary. This means that there does not exist a second, 'initial', boundary on which one would have to specify boundary data. Formally, one would write

$$\Psi[h_{ab}, \Phi, \Sigma] = \sum_{\mathcal{M}} \nu(\mathcal{M}) \int_{\mathcal{M}} \mathcal{D}g\mathcal{D}\Phi\ \mathrm{e}^{-S_{\mathrm{E}}[g_{\mu\nu},\Phi]}. \tag{8.58}$$

The sum over $\mathcal{M}$ expresses the sum over all four-manifolds with measure $\nu(\mathcal{M})$. Since it is known that four-manifolds are not classifiable, this cannot be put into a precise mathematical form. The integral is the quantum-gravitational path integral discussed in Section 2.2.1, where S_{E} denotes the Euclidean Einstein–Hilbert action (2.73).

Except in simple minisuperspace models, the path integral in (8.58) cannot be evaluated exactly. It is therefore usually calculated in a semiclassical ('saddle-point') approximation. Since there exist in general several saddle points, one must address the issue of which contour of integration has to be chosen. Depending on the contour, only some of the saddle points will contribute to the path integral.

The integral over the four-metric in (8.58) splits into integrals over the three-metric, the lapse function, and the shift vector; cf. Section 5.3.4. In a Friedmann model, only the integral over the lapse function remains, and this turns out to be an ordinary integral (cf. Halliwell (1988)),

$$\psi(a,\phi) = \int \mathrm{d}N \int \mathcal{D}a\mathcal{D}\phi\ \mathrm{e}^{-I[a(\tau),\phi(\tau),N]}, \tag{8.59}$$

where we have denoted the Euclidean minisuperspace action by I instead of S_{E}. For notational simplicity, we have denoted the arguments of the wave function by the same letters as for the corresponding functions that are integrated over in the path integral. For the Friedmann model containing a scalar field, the Euclidean action reads

$$I = \frac{1}{2}\int_{\tau_1}^{\tau_2} \mathrm{d}\tau\ N\left[-\frac{a}{N^2}\left(\frac{\mathrm{d}a}{\mathrm{d}\tau}\right)^2 + \frac{a^3}{N^2}\left(\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right)^2 - a + a^3V(\phi)\right], \tag{8.60}$$

where $V(\phi)$ may denote just a mass term, that is, $V(\phi) \propto m^2\phi^2$, or include a self-interaction such as $V \propto \phi^4$.

How is the no-boundary proposal implemented in minisuperspace? The restriction on the class of contours in (8.58) is implemented here by integration over Euclidean paths $a(\tau)$ with the boundary condition $a(0) = 0$; cf. the discussion in Halliwell (1991). This is supposed to implement the idea of integration over regular four-geometries with no 'boundary' at $a = 0$. (The point $a = 0$ has to be viewed like a pole of a sphere, which is completely regular.) One is often interested in discussing quantum-cosmological models in the context of inflationary cosmology. Therefore, in evaluating (8.59), one might restrict oneself to the region where the scalar field ϕ is slowly varying (the 'slow-roll approximation' of inflation). One can then neglect the kinetic term of ϕ and integrate over Euclidean paths with $\phi(\tau) \approx$ constant. For $a^2V < 1$, one gets the following two saddle point actions (Hawking 1984, Halliwell 1991):

$$I_{\pm} = -\frac{1}{3V(\phi)}\left[1 \pm (1 - a^2V(\phi))^{3/2}\right]. \tag{8.61}$$

The action I_- is obtained for a three-sphere that is closed off by less than half of the four-sphere, while the three-sphere is closed off by more than half of the four-sphere in the evaluation of I_+.

There exist various arguments as to which of the two extremal actions is distinguished by the no-boundary proposal. This can, in general, only be decided by a careful discussion of integration contours in the complex N-plane; see, for example, Halliwell and Louko (1990) or Kiefer (1991). For the present purpose, it is sufficient to assume that I_- gives the dominant contribution (Hartle and Hawking 1983). The wave functions $\psi \propto \exp(-I_\pm)$ are WKB solutions of the minisuperspace Wheeler–DeWitt equation (8.25) in the classically forbidden ('Euclidean') region. Taking the standard WKB prefactor into account, one thus has for the no-boundary wave function (choosing I_-) for $a^2V < 1$ the expression

$$\psi_{\rm NB} \propto \left(1 - a^2V(\phi)\right)^{-1/4} \exp\left(\frac{1}{3V(\phi)}\left[1 - (1 - a^2V(\phi))^{3/2}\right]\right). \tag{8.62}$$

Note that the sign in the exponential has been fixed by the proposal—the WKB approximation would also allow a solution of the form $\propto \exp(-\ldots)$. The continuation into the classically allowed region $a^2V > 1$ is obtained through the standard WKB connection formulae to read

$$\psi_{\rm NB} \propto \left(a^2V(\phi) - 1\right)^{-1/4} \exp\left(\frac{1}{3V(\phi)}\right) \cos\left(\frac{(a^2V(\phi) - 1)^{3/2}}{3V(\phi)} - \frac{\pi}{4}\right). \tag{8.63}$$

The no-boundary proposal thus picks out a particular WKB solution in the classically allowed region. It is a real solution and can therefore be interpreted as a superposition of two complex WKB solutions of the form $\exp(\mathrm{i}S)$ and $\exp(-\mathrm{i}S)$.

The above wave functions have been obtained for regions of slowly varying ϕ. This is the context of inflationary cosmology, in which one has an effective (ϕ-dependent) Hubble parameter of the form

$$H^2(\phi) \approx \frac{4\pi V(\phi)}{3m_{\rm P}^2}, \tag{8.64}$$

where, in the simplest case, one has $V(\phi) = m^2\phi^2$. In the units used here ($3m_{\rm P}^2 = 4\pi$), this reads $H^2(\phi) = V(\phi)$. The radius of the four-sphere is $a = H^{-1} = V^{-1/2}$. The geometric picture underlying the no-boundary proposal here is one in which the dominant geometry in the path integral consists of two parts: half of a four-sphere, to which half of de Sitter space is attached. The matching must be made at exactly half of the four-sphere because only there is the extrinsic curvature equal to zero. Only for vanishing extrinsic curvature, $K_{ab} = 0$, is continuity guaranteed; cf. Gibbons and Hartle (1990). The resulting geometry is called the 'Hartle–Hawking instanton' or 'real tunnelling geometry'; see Fig. 8.3.[6] From (8.61), it is clear that the action corresponding to half of the four-sphere is $I = -1/3V(\phi)$. The solutions of the classical field equations are $a(\tau) = H^{-1}\sin(H\tau)$ in the Euclidean regime ($0 \le \tau \le \pi/2H$) and $a(t) = H^{-1}\cosh(Ht)$ in the Lorentzian regime ($t > 0$).

This picture of de Sitter space attached to half of a four-sphere is often referred to as 'quantum creation from nothing' or 'nucleation' from the Euclidean regime into de Sitter space. However, this is somewhat misleading since it is not a process in time, but corresponds to the emergence *of* time (cf. Butterfield and Isham 1999).

[6]This instanton was discussed earlier by Vilenkin (1982).

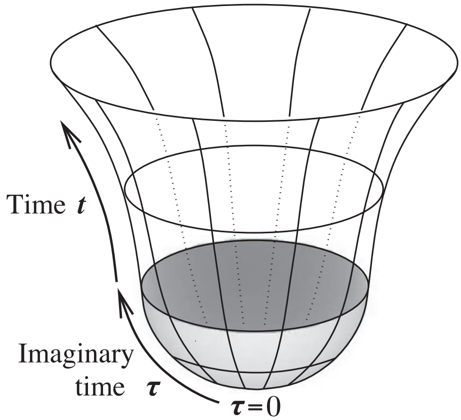

Fig. 8.3 Hartle–Hawking instanton: the dominant contribution to the Euclidean path integral is assumed to be from half of a four-sphere attached to a part of de Sitter space.

Moreover, it is far from clear that this is really the dominant contribution to the path integral: discussions in 2+1 dimensions, where more explicit calculations can be made, indicate that the path integral is dominated by an infinite number of complicated topologies (Carlip 1998). This would cast doubt on the validity of the above saddle-point approximation in the first place.

It is not clear that a real tunnelling geometry as described above exists in all models. In a general situation, one has to look for integration contours in the space of *complex* metrics that render the path integral convergent. Consider the quantum-cosmological example defined by the Wheeler–DeWitt equation (8.28). As mentioned above, the classical solutions are confined to a rectangle centred around the origin in the (a, χ) plane. One can explicitly construct wave packets that follow these classical trajectories; cf. Fig. 8.2. The corresponding quantum states are normalizable in both the a and the χ direction. A general quantum-gravitational (cosmological) path integral depends on two pairs of values for a and χ, called a', a'' and χ', χ'', respectively. Figure 8.4 shows the corresponding (a'', χ'') space; the values for a' and χ' define the origin of the 'lightcones' depicted in the figure by bullets.

The Hartle–Hawking wave function is obtained for $a' = 0 = \chi'$, that is, for the case when the lightcones in Fig. 8.4 shrink to one cone centred at the origin. For this model, the path integral can be evaluated *exactly* (Kiefer 1991). The result shows that there is no contour for the path integration in the complex-metric plane that leads to a wave function which can be used in the construction of wave packets following the classical trajectories: the resulting wave functions either diverge along the 'lightcones' or diverge for large values of a and χ. The states are thus not normalizable, and it is not clear how they should be interpreted. This gives an idea of the problems that result if one goes beyond the semiclassical approximation.

An interesting consequence of the no-boundary proposal occurs if the inhomogeneous modes of Section 8.2 are taken into account. In fact, this proposal selects a distinguished vacuum for de Sitter space—the so-called 'Euclidean' or 'Bunch–Davies'

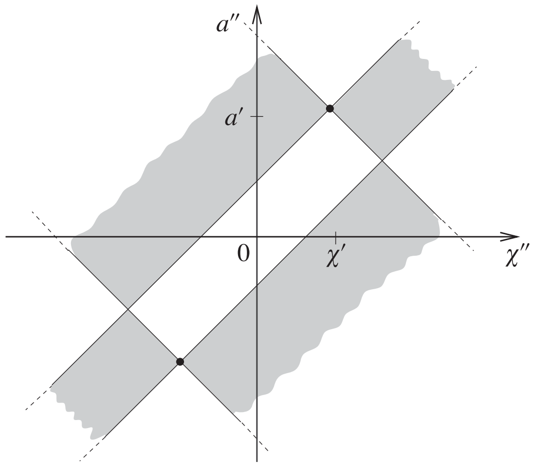

Fig. 8.4 The wave functions obtained from the path integral for the model defined by (8.28) diverge either along the lightcones in minisuperspace or for large values of a and χ. They exhibit oscillatory behaviour in the shaded regions.

vacuum (Laflamme 1987). What is the Euclidean vacuum? In Minkowski space, there exists a distinguished class of equivalent vacua (simply called the 'Minkowski vacuum'), which is invariant under the Poincaré group and therefore the same for all inertial observers. De Sitter space is, like Minkowski space, maximally symmetric: instead of the Poincaré group, it possesses SO(4, 1) (the 'de Sitter group') as its isometry group, which also has 10 parameters. It turns out that, in contrast to Minkowski space, there exists for massive quantum fields a one-parameter family of inequivalent vacua which are invariant under the de Sitter group; see, for example, Birrell and Davies (1982). One of these vacua is distinguished in many respects: it corresponds to the Minkowski vacuum for constant a, and its mode functions are regular on the Euclidean section $t \mapsto \tau = \mathrm{i}t + \pi/2H$. This second property gives it the name 'Euclidean vacuum'.

As in (8.38), one expands the scalar field in its harmonics,

$$\Phi(\mathbf{x}, \tau) - \frac{1}{\sqrt{2}\pi}\phi(\tau) = \sum_{\{n\}} f_n(\tau) Q^n ,$$

but now with respect to the Euclidean time τ. One then calculates the Euclidean action for the modes $\{f_n(\tau)\}$ and imposes the following regularity conditions from the no-boundary proposal:

$$f_n(0) = 0 \ , \ n = 2, 3, \dots \ , \quad \frac{\mathrm{d}f_n}{\mathrm{d}\tau}(0) = 0 \ , \ n = 1.$$

For the wave functions ψ_n satisfying the Schrödinger equations (8.49) this then yields a solution that corresponds to the Euclidean vacuum. The reason is that essentially the same regularity conditions are required for the no-boundary proposal and the Euclidean vacuum. In the Lorentzian section, the Euclidean vacuum for modes with small wavelength (satisfying $\lambda \ll H^{-1}$) is given by the state (8.56); see also Section 10.1. According to the no-boundary proposal, the multipoles thus occur in the Lorentzian regime in their ground state. Because of its high symmetry, the de Sitter-invariant vacuum was assumed to be a natural initial quantum state even before the advent of the no-boundary condition (Starobinsky 1979).

Hawking (1984) suggested that the Euclidean path integral is the true fundamental concept. The fact that a Euclidean metric usually does not have a Lorentzian section therefore does not matter. Only the result—the wave function—counts. If the wave function turns out to be exponentially increasing or decreasing, it describes a classically forbidden region. If it is of oscillatory form, it describes a classically allowed region—this corresponds to the world we live in. Since in general one has to use complex integration contours anyway, it is clear that only the result can have interpretational value, with the formal manipulations playing only a heuristic role.

8.3.3 Tunnelling condition

The no-boundary wave function calculated in the last subsection turned out to be real. This is a consequence of the Euclidean path integral; even if complex metrics contribute, they should do so in complex-conjugate pairs. The wave function (8.63) can be written as a sum of semiclassical components of the form $\exp(\mathrm{i}S)$, each of which gives rise to a semiclassical world in the sense of Section 5.4 (recovery of the Schrödinger equation). These components become independent of each other only after decoherence is taken into account; see Section 10.1. Alternative boundary conditions may give a complex wave function directly, of the form $\exp(\mathrm{i}S)$ in the semiclassical approximation. This is achieved by the 'tunnelling proposal' put forward by Vilenkin; see, for example, Vilenkin (1988, 2003).[7]

The tunnelling proposal is most easily formulated in minisuperspace. In analogy with, for example, the process of α-decay in quantum mechanics, it is proposed that the wave function consists solely of *outgoing* modes. More generally, the proposal states that the wave function consists solely of outgoing modes at singular boundaries of superspace (except for the boundaries corresponding to vanishing three-geometry). In the minisuperspace example above, this is the region of infinite a or ϕ. What does 'outgoing' mean? The answer is clear in quantum mechanics, since there one has a reference phase $\propto \exp(-\mathrm{i}\omega t)$. An outgoing plane wave would then have a wave function $\propto \exp(\mathrm{i}kx)$. But since there is no external time t in quantum cosmology, one can call a wave function 'outgoing' only by definition (Zeh 1988). In fact, the whole concept of tunnelling loses its meaning if an external time is lacking (Conradi 1998).

We have seen in (5.22) that the Wheeler–DeWitt equation possesses a conserved 'Klein–Gordon current', which here takes the form

$$j = \frac{\mathrm{i}}{2}(\psi^* \nabla \psi - \psi \nabla \psi^*) , \quad \nabla j = 0 \tag{8.65}$$

(∇ denotes again the derivatives in minisuperspace). A WKB solution of the form $\psi \approx C \exp(\mathrm{i}S)$ leads to

[7] Like the no-boundary proposal, this usually refers to a closed three-space Σ. A treatment of 'tunnelling' into a universe with an open Σ is presented in Zel'dovich and Starobinsky (1984).

$$j \approx -|C|^2 \nabla S. \tag{8.66}$$

The tunnelling proposal states that this current should point outwards at large a and ϕ (provided, of course, that ψ is of WKB form there). If ψ were real (as is the case in the no-boundary proposal), the current would vanish.

In the above minisuperspace model, we have seen that the eikonal $S(a,\phi)$, which is a solution of the Hamilton–Jacobi equation, is given by the expression (cf. (8.63))

$$S(a,\phi) = \frac{(a^2V(\phi)-1)^{3/2}}{3V(\phi)}. \tag{8.67}$$

We would thus have to take the solution $\propto \exp(-\mathrm{i}S)$, since then j would, according to (8.66), become positive and point outwards for large a and ϕ. For $a^2V > 1$, the tunnelling wave function then reads

$$\psi_{\mathrm{T}} \propto (a^2V(\phi)-1)^{-1/4} \exp\left(-\frac{1}{3V(\phi)}\right) \exp\left(-\frac{\mathrm{i}}{3V(\phi)}(a^2V(\phi)-1)^{3/2}\right), \tag{8.68}$$

while for $a^2V < 1$ (the classically forbidden region), one has

$$\psi_{\mathrm{T}} \propto (1-a^2V(\phi))^{-1/4} \exp\left(-\frac{1}{3V(\phi)}\left(1-(1-a^2V(\phi))^{3/2}\right)\right). \tag{8.69}$$

If the inhomogeneous modes are taken into account as earlier, the tunnelling proposal also picks out the Euclidean vacuum.

8.3.4 Comparison of no-boundary and tunnelling wave functions

An important difference between the no-boundary and the tunnelling condition is the following: whereas the tunnelling condition is imposed in the oscillatory regime of the wave function, the no-boundary condition is implemented in the Euclidean regime; the oscillatory part of the wave function is then found by a matching procedure. In the above example, this leads to the crucial difference between ψ_{T} and ψ_{NB} that, besides having a real versus a complex wave function, ψ_{T} contains a factor $\exp(-1/3V)$, whereas ψ_{NB} has $\exp(1/3V)$. Under the assumption that our branch of the wave function is in some sense dominant, these results for the wave function have been used to calculate the probability for the occurrence of an inflationary phase. More precisely, the question of whether the wave function favours large values of ϕ (as would be needed for inflation) or small values has been investigated. It is clear from the above results that ψ_{T} favours large ϕ over small ϕ and therefore seems to predict inflation, whereas ψ_{NB} prefers small ϕ and therefore seems to predict no inflation. However, the assumption of the slow-roll approximation (needed for inflation) is in contradiction to sharp probability peaks; see Barvinsky (2001) and the references therein. The reason is that this approximation demands the ϕ-derivatives to be small. A possible way out of this dilemma is to take into account inhomogeneities (the higher multipoles of Section 8.2) and to proceed to the one-loop approximation of the wave function. This ensures normalizability of the wave function provided that certain restrictions on the particle content of the theory are fulfilled. Let us take a closer look at this.

The wave function in the Euclidean one-loop approximation is given by (Barvinsky and Kamenshchik 1990)

$$\Psi_{\mathrm{T,NB}} = \exp(\pm I - W), \tag{8.70}$$

where the T and NB refer to 'tunnelling' and 'no-boundary', respectively, I is the classical Euclidean action,[8] and W is the one-loop correction to the effective action,

$$W = \frac{1}{2}\mathrm{tr}\ln\frac{F}{\mu^2}. \tag{8.71}$$

Here, F represents the second-order differential operator that is obtained by taking the second variation of the action with respect to the fields, while μ is a renormalization mass parameter. It can be shown that, after analytic continuation into the Lorentzian regime, the wave function assumes the form

$$\Psi_{\mathrm{T,NB}} = \left(\frac{1}{|\det u|^{1/2}}\right)^R \times \exp\left(\pm I + \mathrm{i}S + \frac{1}{2}\mathrm{i}f^T(Dv)v^{-1}f\right). \tag{8.72}$$

Here, S is the minisuperspace part of the Lorentzian classical action; f denotes the amplitudes of the inhomogeneities of the geometry and matter (Section 8.2), u the solutions of the linearized equations for all the modes, and v those solutions referring to the inhomogeneous modes; D is a first-order differential operator (the Wronskian related to the operator F), and the superscript R denotes the renormalization of the infinite product of basis functions.

Information concerning the normalizability of the wave function of the universe (8.72) can be obtained from the diagonal elements of the density matrix

$$\hat{\rho} = \mathrm{tr}_f|\Psi\rangle\langle\Psi|. \tag{8.73}$$

It was shown in Barvinsky and Kamenshchik (1990) that the diagonal elements (denoted by $\rho(\phi)$) can be expressed as

$$\rho(\phi) \sim \exp(\pm I - \Gamma_{1-\mathrm{loop}}), \tag{8.74}$$

where $\Gamma_{1-\mathrm{loop}}$ is the one-loop correction to the effective action calculated on the closed compact 'Hartle–Hawking instanton' (Fig. 8.3). This quantity is conveniently calculated by the zeta-regularization technique (Birrell and Davies 1982), which allows $\Gamma_{1-\mathrm{loop}}$ to be represented as

$$\Gamma_{1-\mathrm{loop}} = \tfrac{1}{2}\,\zeta'(0) - \tfrac{1}{2}\,\zeta(0)\ln(\mu^2 a^2), \tag{8.75}$$

where $\zeta(s)$ is the generalized Riemann zeta function, and a is the radius of the instanton. In the limit $\phi \to \infty$ or, equivalently, $a \to 0$, the expression (8.74) reduces to

$$\rho(\phi) \sim \exp(\pm I)\phi^{-Z-2}, \tag{8.76}$$

[8]In the above example, $I = -(3V)^{-1}[1 - (1 - a^2V)^{3/2}]$, and $\exp(+I)$ results from the tunnelling condition, while $\exp(-I)$ results from the no-boundary condition.

where Z is the anomalous scaling of the theory, expressed in terms of $\zeta(0)$ for all the fields included in the model. The requirement of normalizability imposes the following restriction on Z:

$$Z > -1, \tag{8.77}$$

which is obtained from the requirement that the integral

$$\int^{\infty} \mathrm{d}\phi \; \rho(\phi)$$

converge at $\phi \to \infty$. In view of this condition, it turns out that SUSY models seem to be preferred (Kamenshchik 1990).

Having obtained the one-loop order, one can investigate whether the wave function is peaked at values of the scalar field favourable for inflation (Barvinsky and Kamenshchik 1998; Barvinsky 2001). It turns out that this works only if the field ϕ is coupled non-minimally: one must have a coupling $-\xi R\phi^2/2$ in the action with $\xi < 0$ and $|\xi| \gg 1$. One can then get a probability peak at a value of ϕ corresponding to an energy scale needed for inflation (essentially the GUT scale). It seems that, again, only the tunnelling wave function can fulfil this condition, although the last word on this has not been spoken. One obtains from this consideration also a restriction on the particle content of the theory in order to obtain $\Delta\phi/\phi \sim \Delta T/T$ in accordance with the observational constraint $\Delta T/T \sim 10^{-5}$ (the temperature anisotropy in the cosmic microwave background).

A detailed investigation was also made for gravity coupled non-minimally to the Standard Model Higgs field (Barvinsky *et al.* 2010), a model that can be used to take the Higgs field as the inflaton for an inflationary scenario. More precisely, the Lagrangian of the graviton–inflaton sector reads

$$\begin{aligned} \boldsymbol{L}(g_{\mu\nu}, \varPhi) &= \frac{1}{2}\left(M_{\mathrm{P}}^2 + \xi|\varPhi|^2\right) R - \frac{1}{2}|\nabla\varPhi|^2 - V(|\Phi|), \\ V(|\Phi|) &= \frac{\lambda}{4}(|\varPhi|^2 - v^2)^2, \quad |\varPhi|^2 = \varPhi^\dagger\varPhi, \end{aligned}$$

where $\varPhi$ denotes the Standard Model Higgs field, whose expectation value plays the role of an inflaton here and which is assumed to possess a strong non-minimal curvature coupling with $\xi \gg 1$; M_{P} is the reduced Planck mass (1.7). The result for the diagonal elements of the density matrix of the tunnelling wave function turns out to be

$$\rho_{\mathrm{tunnel}}(\varphi) = \exp\left(-\frac{24\pi^2 M_{\mathrm{P}}^4}{\hat{V}(\varphi)}\right), \tag{8.78}$$

where $\varphi := (\Phi^\dagger\Phi)^{1/2}$, and $\hat{V}(\varphi)$ denotes the effective Higgs potential in the Einstein frame of the action. It is clear from (8.78) that the probability to have a sufficiently large φ is highest at the maximum of the effective potential. For the lowest possible Higgs mass (below which the potential develops an instability in this model), the potential is shown in Fig. 8.5. One recognizes that the potential is maximal at a field value that is large enough to be suitable for the onset of inflation.[9]

[9]Decoherence also sets in at the onset of inflation (see Section 10.2); it is thus justified to concentrate here on the probabilities and to neglect interference terms.

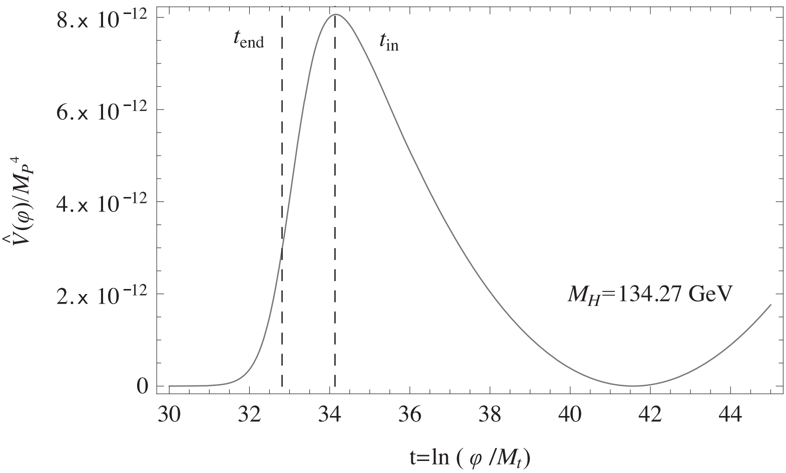

Fig. 8.5 Effective potential for the Higgs field coupled non-minimally to gravity at the lowest possible Higgs mass in the stability regime. The parameter t denotes a dimensionless time (given by the logarithm of the ratio of the Higgs field to the top-quark mass) that increases from right to left. The dashed lines denote a possible inflationary domain between $t_{\rm in}$ and $t_{\rm end}$. From Barvinsky *et al.* (2010). © 2010 by the American Physical Society.

We mention, finally, that there is a difference concerning the factor ordering (Kontoleon and Wiltshire 1996): it seems that the tunnelling condition can only be consistently defined for particular factor orderings, in contrast to the no-boundary condition.

8.3.5 Symmetric initial condition

This condition was proposed by Conradi and Zeh (1991); see also Conradi (1992). For the wave-packet solutions of Section 8.1, we had to demand that ψ goes to zero as $\alpha \to \infty$. Otherwise, the packet would not reflect the behaviour of a classically recollapsing universe. But what about the behaviour at $\alpha \to -\infty$ ($a \to 0$)? Consider again the model of a massive scalar field in a Friedmann universe given by the Wheeler–DeWitt equation (8.25). The potential term vanishes in the limit $\alpha \to -\infty$, so the solutions that are exponentially decreasing for large α become constant in this limit. With regard to finding normalizable solutions, it would be ideal if there were a reflecting potential also at $\alpha \to -\infty$. One can add for this purpose, in an ad hoc manner, a repulsive (negative) potential that would be of relevance only in the Planck regime. The application of loop quantum gravity to cosmology, as discussed in the next section, can lead to the occurrence of such a respulsive potential. It is also imaginable that it results from a unification of interactions. One can, for example, choose the 'Planck potential'

$$V_{\rm P}(\alpha) = -C^2 {\rm e}^{-2\alpha}, \tag{8.79}$$

where C is a real constant. Neglecting, as in the previous subsections, the ϕ-derivatives (corresponding to the slow-roll approximation), one can thereby *select* a solution to the Wheeler–DeWitt equation ((8.25) supplemented by $V_{\rm P}$) that decreases exponentially as $\alpha \to -\infty$. This also implements DeWitt's boundary condition that $\psi \to 0$ for $\alpha \to -\infty$ (Section 8.3.1).

The 'symmetric initial condition' (SIC) now states that in the limit $\alpha \to -\infty$, the full wave function depends only on α; cf. Conradi and Zeh (1991). In other words, it is a particular superposition (not an ensemble) of all excited states of Φ and the three-metric; that is, these degrees of freedom are completely absent in the wave function. This is analogous to the symmetric vacuum state in field theory before the symmetry breaking into the 'false' vacuum (Zeh 2007). In both cases, the actual symmetry breaking will occur through decoherence (Section 10.2). The resulting wave function coincides approximately with the no-boundary wave function. As in that case, the higher multipoles enter the semiclassical Friedmann regime in their ground state. The SIC is also well suited for a discussion of the arrow of time and the dynamical origin of irreversibility; cf. Section 10.3.

8.4 Loop quantum cosmology

8.4.1 Classical variables

The cosmological applications of quantum gravity discussed above make use of geometrodynamical variables. However, quantum cosmology can also be treated using the loop variables discussed in Chapter 6. The ensuing framework of loop quantum cosmology was introduced by Martin Bojowald; see Bojowald (2011) and Ashtekar and Singh (2011) for a detailed review and references. The mathematical structure of loop quantum cosmology is presented in Ashtekar *et al.* (2003); a brief introduction with emphasis on the phenomenological aspects has been given by Sakellariadou (2010). As in the minisuperspace approach discussed above, loop quantum cosmology is constructed via a truncation of the classical phase space of GR to spatially homogeneous situations, which is then quantized by using the methods and results of loop quantum gravity (Chapter 6). Features such as the quantization of geometric operators are thereby transferred to the truncated models. In the present section, we restrict ourselves to the simplest case of Friedmann universes; anisotropic models and inhomogeneous situations can also be addressed (Bojowald 2011). As in quantum geometrodynamics, loop quantum cosmology is not derived from the full theory (here, loop quantum gravity); the hope is, again, that the quantization of the classically truncated models will reflect some of the true features of quantum cosmology.

We consider the model of a Friedmann universe containing a scalar field; cf. Section 8.1.2. Instead of the original variables a and p_a, we shall use new canonical variables which result from the truncation of the general canonical variables of the holonomy and triad to the homogeneous, isotropic model. How this truncation is performed in a mathematically clean way is shown in detail in Bojowald (2011) and the references therein. From the triad one obtains the single variable $\tilde{p}$, while the holonomy leads to the single variable $\tilde{c}$, where $A_a^i =: \tilde{c}\delta_a^i$. What is their explicit form? We shall assume in the following a Friedmann universe with a finite spatial volume V_0 and allow it to be either positively curved ($k = 1$) or flat ($k = 0$). The new variables are then obtained from the ones in Section 8.1.2 by

$$|\tilde{p}| = a^2 \ , \quad \tilde{c} = k + \beta\dot{a}, \tag{8.80}$$

where, from (8.9), we have

$$\dot{a} = -\frac{4\pi G}{3V_0}\frac{Np_a}{a},$$

and β is the Barbero–Immirzi parameter introduced in Section 4.3.1. The Poisson bracket between the new variables reads

$$\{\tilde{c}, \tilde{p}\} = \frac{8\pi G\beta}{3V_0}.$$

It is convenient to absorb the volume V_0 into the canonical variables by the substitution

$$\tilde{p} = V_0^{-2/3}p\ , \quad \tilde{c} = V_0^{-1/3}c,$$

leading to

$$\{c, p\} = \frac{8\pi G\beta}{3}. \tag{8.81}$$

We note that p has the physical dimensions of a length squared, while c is dimensionless. The sign of p reflects the orientation of the triad. Both orientations are thus present in the formalism. The classical singularity, which lies at $p = 0$, is thus not a boundary; instead, it is an interior point of the configuration space. This will have important consequences.

The Hamiltonian constraint (8.10) can easily be rewritten in terms of the new variables. If the lapse function is chosen as $N = 1$, the constraint reads (using the identity $k^2 = k$)

$$H = -\frac{3}{8\pi G}\left(\frac{(c-k)^2}{\beta^2} + k^2\right)\sqrt{|p|} + H_{\mathrm{m}} \approx 0, \tag{8.82}$$

where, in the case of the scalar field,

$$H_{\mathrm{m}} = \frac{1}{2}\left(|p|^{-3/2}p_\phi^2 + |p|^{3/2}\mathcal{V}(\phi)\right). \tag{8.83}$$

The starting point is now set for the quantization.

8.4.2 Quantization

In accordance with the spirit of full loop quantum gravity, one does not quantize c directly, but the related holonomy. The reason is that in loop quantum gravity A_a^i is not represented by a local operator, and so the same holds for its truncated version c. This feature also lies at the heart of the discreteness of loop quantum gravity discussed in Section 6.2.

In the case of the minisuperspace of a Friedmann–Lemaître model, one has to consider special holonomies that are constructed from isotropic connections. This introduces a parameter length μ (chosen to be dimensionless) into the formalism, which captures information about the edges and the spin labels of the spin network (μ is not a physical length). The fact that such a parameter appears is connected with the fact that the reduced model still possesses a background structure, in spite of the background independence of the full theory. This is because the spatial metric is unique up to the scale factor, which thus introduces a conformal space as a background.

The fact that c itself is not turned into an operator, but only the holonomy constructed from it, has an important consequence: loop quantum cosmology requires a 'new quantum mechanics', that is, it employs a representation of the canonical variables that is inequivalent to the standard Schrödinger representation. We shall see the reason for this more clearly below.

Instead of turning the classical Poisson-bracket relation (8.81) into a commutator acting on the standard Hilbert space, in the 'new quantum mechanics' of loop quantum cosmology one makes use of the *Bohr compactification*[10] of $\mathbb{R}$ as the appropriate framework for this 'holonomy representation'; cf. Ashtekar *et al.* (2003). One starts with the algebra of functions of the form

$$f(c) = \sum_{\mu} f_\mu \mathrm{e}^{\mathrm{i}\mu c/2} \tag{8.84}$$

(which are called almost-periodic functions), where the sum runs over a countable subset of $\mathbb{R}$. The motivation for this is that the reduction procedure from the holonomies of the full theory leads to the factor $\mathrm{e}^{\mathrm{i}\mu c/2}$. This algebra is isomorphic to the Bohr compactification of $\mathbb{R}$, $\bar{\mathbb{R}}_{\mathrm{Bohr}}$, which is a compact group and contains $\mathbb{R}$ densely. It can be obtained as the dual group of the real line endowed with the discrete topology. Representations of $\bar{\mathbb{R}}_{\mathrm{Bohr}}$ are labelled by real numbers μ and are given by[11]

$$\rho_\mu : \bar{\mathbb{R}}_{\mathrm{Bohr}} \to \mathbb{C}\ , \quad c \mapsto \mathrm{e}^{\mathrm{i}\mu c}.$$

If it had been possible to quantize c directly, we would have chosen the space of square-integrable functions, $L^2(\mathbb{R}\,\mathrm{d}c)$, as the Hilbert space. In the Bohr compactification scheme, we choose instead the space $L^2(\bar{\mathbb{R}}_{\mathrm{Bohr}}, \mathrm{d}\mu(c))$, where $\mathrm{d}\mu(c)$ is the Haar measure, defined by

$$\int_{\bar{\mathbb{R}}_{\mathrm{Bohr}}} \mathrm{d}\mu(c)\ f(c) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{d}c\ f(c).$$

The Hilbert space of loop quantum cosmology is thus the space of all square-integrable functions on a compactification of the real line. The basis states are chosen to be

$$\langle c|\mu\rangle = \mathrm{e}^{\mathrm{i}\mu c/2}, \tag{8.85}$$

which obey

$$\langle \mu_1|\mu_2\rangle = \delta_{\mu_1,\mu_2}. \tag{8.86}$$

In the standard Schrödinger representation, one would have the delta function $\delta(\mu_1 - \mu_2)$ on the right-hand side here. The occurrence of the Kronecker symbol means that

[10]This concept is named after Harald August Bohr (1887–1951), the younger brother of Niels Bohr, who contributed much to the theory of almost-periodic functions.

[11]To emphasize the discrete nature of $\bar{\mathbb{R}}_{\mathrm{Bohr}}$, one also talks in this context of the *polymer representation* of the abstract Weyl algebra.

the basis elements $|\mu\rangle$, $\mu \in \mathbb{R}$, are uncountable and that the Hilbert space is therefore non-separable.

How is p to be quantized? Since, according to (8.81), it is conjugate to c, one represents it by a derivative operator,

$$\hat{p} = -\mathrm{i}\frac{8\pi\beta l_{\mathrm{P}}^2}{3}\frac{\mathrm{d}}{\mathrm{d}c}, \tag{8.87}$$

leading to[12]

$$\hat{p}|\mu\rangle = \frac{4\pi\beta l_{\mathrm{P}}^2}{3}\mu|\mu\rangle =: p_\mu|\mu\rangle. \tag{8.88}$$

The basis states are thus eigenstates of the 'flux operator' $\hat{p}$. This is possible because of the orientation freedom for the triads, allowing $\mu \in \mathbb{R}$. However, the spectrum of $\hat{p}$ is 'discrete' in the sense that its eigenstates $|\mu\rangle$ are normalizable (this is possible because the Hilbert space is non-separable). The kinematical discreteness of the full theory thus survives in the truncated version. We note also the relation

$$\widehat{\mathrm{e}^{\mathrm{i}\mu' c/2}}|\mu\rangle = |\mu + \mu'\rangle, \tag{8.89}$$

from which one gets

$$\langle\mu|\widehat{\mathrm{e}^{\mathrm{i}\mu' c/2}}|\mu\rangle = \delta_{0,\mu'}.$$

From this, one recognizes that the operator $\widehat{\mathrm{e}^{\mathrm{i}\mu' c/2}}$ is not (weakly) continuous in μ', since a vector $|\mu\rangle$ is mapped ('translated') for arbitrarily small μ' to an orthogonal state $|\mu+\mu'\rangle$. This is the reason why the Stone–von Neumann theorem of quantum mechanics (which demands weak continuity of the one-parameter groups $U(\lambda)$ and $V(\mu)$ satisfying the Weyl commutation relations) does not apply here and why one thus obtains a representation which is inequivalent to the standard Schrödinger representation.

We shall now turn to a discussion of the quantum Hamiltonian constraint for the Friedmann universe with a scalar field. Consider first the matter Hamiltonian (8.83). The major problem is to construct a well-defined operator for the classically divergent (in the limit $a \to 0$) expression $|p|^{-3/2} \propto a^{-3}$. As can be seen from (8.88), the operator $\hat{p}$ possesses a discrete spectrum containing zero and is thus not invertible. In order to deal with this situation, one proceeds as in the full theory and makes use of a 'Poisson-bracket trick', in which one uses the identity (4.137) and transforms the Poisson bracket occurring therein into a commutator. Here, instead of $|p|^{-3/2}$, one directly uses the function

$$d(p) = \frac{1}{3\pi\beta G}\sum_{i=1}^{3}\mathrm{tr}(\tau_i U_i\{U_i^{-1}, \sqrt{V}\})^6, \tag{8.90}$$

where the U_i denote the holonomies of the isotropic connections, and $V = |p|^{3/2}$ is the volume. For large p one has $d(p) \sim |p|^{-3/2}$, as required. Turning the holonomies into

[12] The difference in the numerical factors appearing in some of the cited references is due to the fact that in much of the literature on loop quantum cosmology, $8\pi l_{\mathrm{P}}^2$ is used instead of l_{P}^2.

multiplication operators and the Poisson bracket into a commutator, one indeed finds a densely defined bounded operator. Its spectrum is given by

$$\hat{d}(p)|\mu\rangle = \left(\frac{1}{2\pi\beta l_{\mathrm{P}}^2}(\sqrt{V_{\mu+1}} - \sqrt{V_{\mu-1}})\right)^6 |\mu\rangle, \tag{8.91}$$

where

$$V_\mu := |p_\mu|^{3/2} = \left(\frac{4\pi\beta l_{\mathrm{P}}^2}{3}|\mu|\right)^{3/2} ; \tag{8.92}$$

cf. (8.88). There are, of course, ambiguities in defining such an operator. A more general class of functions, $d(p)_{j,l}$, is obtained by introducing the parameters $j \in \frac{1}{2}\mathbb{N}$ and l, $0 < l < 1$, where j arises from the freedom to use different representations of SU(2), and l from the classical freedom in writing $V^{-1} = (V^{l-1})^{1/(1-l)}$. This then leads to an effective matter Hamiltonian

$$H_{\mathrm{m}}^{(\mathrm{eff})} = \frac{1}{2}\left(d(p)_{j,l}p_\phi^2 + |p|^{3/2}\mathcal{V}(\phi)\right). \tag{8.93}$$

The new term $d(p)_{j,l}$ appearing here has various consequences. It gives rise to modified densities in the effective cosmological equations (e.g. the Friedmann equation and the Raychaudhuri equation) and to a modified 'damping term' in the effective Klein–Gordon equation for the scalar field. This leads to qualitative changes at small a, since d does not go to infinity as a approaches zero. One finds an effective *repulsion* or *bounce*, which can potentially prevent the big bang; see Section 8.5. Effectively, this corresponds to the presence of a Planck potential as introduced by Conradi and Zeh (1991); cf. Section 8.3.5. These effects arise exclusively from the gravitational degrees of freedom.

The new term can also enhance the expansion of the universe at small scales, providing a possible mechanism for inflation from pure quantum-gravitational effects. If inhomogeneities are taken into account, it is also conceivable that observable effects would be found in the anisotropy spectrum of the cosmic microwave background (see below).

What about the gravitational part of the Hamiltonian constraint operator? It contains the term c^2 and thus is not an almost-periodic function. One may use instead, for example, a function proportional to $\delta^{-2}\sin^2\delta c$, where δ labels a quantization ambiguity (there are many more such ambiguities). This reproduces c^2 in the limit for small c, which is why the classical limit follows only for small c. The situation here is thus much more involved than for the matter Hamiltonian, where one just has to address modified densities. Expanding the general solution of the full constraint in terms of volume eigenstates,

$$|\psi\rangle = \sum_\mu \psi_\mu|\mu\rangle, \tag{8.94}$$

one arrives at the following *difference equation* for the coefficients of this expansion:[13]

[13]Restriction is made to a flat Friedmann universe.

$$\begin{aligned}&(V_{\mu+5\delta}-V_{\mu+3\delta})\psi_{\mu+4\delta}(\phi)-2(V_{\mu+\delta}-V_{\mu-\delta})\psi_\mu(\phi)\\&+(V_{\mu-3\delta}-V_{\mu-5\delta})\psi_{\mu-4\delta}(\phi)=-\frac{128\pi^2G\beta^2\delta^3l_{\mathrm{P}}^2}{3}\hat{H}_{\mathrm{m}}(\phi)\psi_\mu(\phi),\end{aligned} \tag{8.95}$$

where V_μ is defined in (8.92). This is an evolution equation with respect to a discrete *intrinsic time* μ (Bojowald 2001*a*). Note, however, that the state $|\psi\rangle$ is not a volume eigenstate. Assuming that $|\psi\rangle$ is sufficiently smooth, one can recover, for large scales, the minisuperspace Wheeler–DeWitt equation with a particular factor ordering (Bojowald 2001*b*). This provides the bridge to the standard formalism of quantum cosmology discussed in Section 8.1. We remark that it is often more convenient to formulate the difference equation (8.95) in terms of the volume instead of μ (Ashtekar and Singh 2011).

Further developments in loop quantum cosmology have dealt with the extension of the formalism to anisotropic and inhomogeneous cosmological models (Bojowald 2011). A comparison of homogeneous and inhomogeneous configurations can be made by using simplified lattice models that are closer in spirit to full loop quantum gravity than the Friedmann models discussed above. Taking inhomogeneities into account in the form of perturbations also allows one to calculate potentially observable effects. One example deals with corrections to the CMB anisotropy spectrum (Bojowald *et al.* 2011). In contrast to the case of Wheeler–DeWitt quantum cosmology (see the end of Section 5.4), the result is an enhancement of the power at large scales.

Results similar to those from loop quantum cosmology can also be obtained from geometrodynamical variables if one employs the Bohr compactification scheme for them, at least for the isotropic case discussed here (Husain and Winkler 2004). Whether this also holds for anisotropic models is not clear. A comparison between quantum cosmology obtained from the Wheeler–DeWitt equation and loop quantum cosmology can be found in the form of a debate in Bojowald *et al.* (2010).

8.5 On singularity avoidance

One of the main goals of any quantum theory of gravity is to cure the singularities prevalent in the classical theory. It is thus an important question whether the approaches discussed in this chapter are able to achieve this for cosmology.

In classical GR, there exist mathematically precise definitions of singularities, as well as rigorous theorems for their occurrence (Hawking and Penrose 1996). So far, this level of rigour does not exist in quantum gravity. One thus has to invoke heuristic ideas and criteria to study the fate of singularities in quantum gravity (Kiefer 2010). One sufficient (but by no means necessary) criterion can be traced back to the pioneering paper on canonical gravity by DeWitt (1967*a*). He suggested that the wave function should vanish in the region of configuation space that corresponds to the classical singularity; see (8.57). Another criterion was suggested in Dąbrowski *et al.* (2006). If a wave packet necessarily spreads when approaching the classical singularity, the semiclassical approximation breaks down and the classical space–time fades away. One can then no longer apply the classical singularity theorems. This second criterion is, of course, a weaker criterion, because no statement is made about singularities in the full quantum theory.

It must be emphasized that the vanishing of the wave function at the classical singularity is only a sufficient, but not a necessary criterion. Let us consider, for example, the solution of the Dirac equation for the ground state of a hydrogen-like atom. In standard notation, the solution reads

$$\psi_0(r) \propto (2mZ\alpha r)^{\sqrt{1-Z^2\alpha^2}-1} \mathrm{e}^{-mZ\alpha r} \overset{r\to 0}{\longrightarrow} \infty$$

and thus diverges at the origin. However, $\int \mathrm{d}r\, r^2|\psi_0|^2$ remains finite, so the important quantity is the measure in the inner product; the probability density goes to zero. In the case of the Wheeler–DeWitt equation, no general consensus on the inner product exists (Section 5.2.2). In minisuperspace models, however, one can choose the standard $\mathcal{L}^2$ inner product. Taking, for example, the solution (8.21) for $k = 0$, one finds a divergence as $a \to 0$,

$$\Psi \propto K_0(a^2/2) \overset{a\to 0}{\longrightarrow} c \ln a,$$

but nonetheless the integral $\int \mathrm{d}a\, \mathrm{d}\phi \sqrt{|G|}\, |\psi(a, \phi)|^2$, where $G = a^2$ is the determinant of the DeWitt metric, remains finite. If this inner product were accepted for the Wheeler–DeWitt equation, this solution would be classified as singularity-free.

Let us consider some examples. The first one is a model with 'phantom fields'. These are fields with negative kinetic energy, which are certainly very exotic but which cannot yet be excluded as the origin of dark energy. Classically, the phantom dynamics develops a big-rip singularity; that is, the energy density ρ and the pressure p diverge as a goes to infinity *in a finite time*. The corresponding quantum model was discussed in Dąbrowski *et al.* (2006). It was found that wave-packet solutions of the Wheeler–DeWitt equation necessarily disperse when approaching the region of the classical big-rip singularity. According to the second of the above criteria, this means that the singularity is avoided in the quantum theory. Quantum effects can thus become important for a large universe, too.

The second example is a model that classically exhibits a big-brake singularity. This is a singularity at which the pressure diverges and where the universe can come to an abrupt halt (infinite deceleration). Such a singularity can be realized by an equation of state of the form $p = A/\rho$, $A > 0$, which describes an 'anti-Chaplygin gas'. For a Friedmann universe with scale factor $a(t) \equiv \mathrm{e}^{\alpha(t)}$ and a scalar field $\phi(t)$, this equation of state can be implemented by choosing the following potential for the scalar field:

$$V(\phi) = V_0 \left(\sinh\left(\sqrt{24\pi G}|\phi|\right) - \frac{1}{\sinh\left(\sqrt{24\pi G}|\phi|\right)} \right), \quad V_0 = \sqrt{A/4}.$$

The quantum version of this model was discussed in Kamenshchik *et al.* (2007). It was shown there that all normalizable solutions of the Wheeler–DeWitt equation have the following form in the vicinity of the classical singularity:

$$\Psi(\alpha, \phi) = \sum_{k=1}^{\infty} A(k) k^{-3/2} K_0 \left(\frac{1}{\sqrt{6}} \frac{V_\alpha}{\hbar^2 k \kappa} \right) \times \left(2\frac{V_\alpha}{k}|\phi| \right) e^{-V_\alpha/k|\phi|} L^1_{k-1} \left(2\frac{V_\alpha}{k}|\phi| \right),$$

where K_0 is a Bessel function, the L^1_{k-1} are Laguerre polynomials, and $V_\alpha \equiv \tilde{V}_0 e^{6\alpha}$. These all *vanish* at the classical singularity. This model therefore implements the first criterion above: the vanishing of the wave function in the spirit of DeWitt (1967*a*). Interestingly, these solutions also implement the avoidance of the big-bang singularity; that is, $\Psi \to 0$ for $\alpha \to -\infty$.

The big-brake model is part of a larger class of models that are described by an equation of state of the form

$$p = -\frac{A}{\rho^\beta},$$

with general real parameters A and β. This is called a 'generalized Chaplygin gas'. The quantum version of these models is discussed in Bouhmadi-López *et al.* (2009). One example is the 'big-freeze singularity', where both the Hubble parameter H and its time derivative $\dot{H}$ blow up in the past at a finite value of the scale factor. The big-freeze singularity occurs in the past at a minimal scale factor $a_{\min} > 0$; there are thus no classical solutions in the limit $\alpha \to -\infty$. For this reason, one has to demand that the wave function goes to zero in the classically forbidden region, $\Psi \stackrel{\alpha \to -\infty}{\longrightarrow} 0$, because otherwise one would not obtain the correct classical limit. The class of solutions to the Wheeler–DeWitt equation then reads (Bouhmadi-López *et al.* 2009)

$$\Psi_k(\alpha, \phi) \propto \sqrt{|\phi|} J_\nu(k|\phi|) \left[b_1 \exp\left(\mathrm{i}\sqrt{\frac{3}{4\pi G}}\alpha \right) + b_2 \exp\left(-\mathrm{i}\sqrt{\frac{3}{4\pi G}}\alpha \right) \right],$$

where $J_\nu(k|\phi|)$ is a Bessel function, and ν is a function of α. These solutions obey DeWitt's boundary condition at the singularity, $\Psi_k(0, 0) = 0$. The same also holds for the other cases discussed in Bouhmadi-López *et al.* (2009). According to the first criterion above, the classical singularities are avoided in the quantum theory.

The vanishing of the wave function at the classical singularity is also used in other applications of the Wheeler–DeWitt equation. An especially interesting example is the supersymmetric quantum cosmological 'billiards' discussed in Kleinschmidt *et al.* (2009): in 11-dimensional supergravity (see Section 2.3), one can employ near a space-like singularity a description based on the Kac–Moody group E_{10} and address the corresponding Wheeler–DeWitt equation. It was found there, too, that $\Psi \to 0$ near the singularity. DeWitt's criterion of singularity avoidance may thus be applicable in a wide range of quantum cosmological models.

Let us now investigate the situation in loop quantum cosmology (Bojowald 2011, Ashtekar and Singh 2011). One of its main new features is the presence of a difference equation instead of a differential equation at the fundamental level (cf. (8.95)). This also presents the role of the classical singularity from a new angle. As emphasized in Section 8.4, there exist two regions in the formalism, differing in the sign of p_μ (which is equal to the sign of μ). These two regions are separated by degenerate geometries at $\mu = 0$. It now turns out that, at least for certain factor orderings, one can evolve the wave functions ψ_μ through $\mu = 0$ from one region into the other, 'jumping over' the singularity. In this sense, the classical singularity is avoided in the quantum theory.

The discrete nature of the difference equation becomes noticeable only when approaching Planck-size universes. For large a, the difference equation becomes indistinguishable from the Wheeler–DeWitt equation, and thus the above heuristic criteria can be applied again. In the case of the classical big-brake singularity, for example, one can find quantum avoidance in the same way as with the Wheeler–DeWitt equation. The situation is, of course, different for the big-bang (and big-crunch) singularities.

In a few of the papers reviewed here, for example Ashtekar and Singh (2011), the question of singularity avoidance has been discussed in quantitative detail for minisuperspace models with a massless scalar field. The presence of this field is needed to serve as an appropriate internal time with respect to which a Hilbert space and a notion of Dirac observables can be introduced. In the context of these models, one finds singularity avoidance and also good semiclassical behaviour for large universes. The quantum dynamics leads to a 'quantum bounce' that corresponds to the presence of a new repulsive potential in the effective Friedmann–Lemaître equations. In the special case of a flat universe, one has

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3}\left(1 - \frac{\rho}{\rho_{\rm c}}\right),$$

where the last term is responsible for the bounce: at the critical density $\rho_{\rm c}$, the expansion of the universe is zero. The critical density turns out to be

$$\rho_{\rm c} = \frac{\sqrt{3}}{32\pi^2\beta^3 G^2\hbar} = \frac{\sqrt{3}\rho_{\rm P}}{32\pi^2\beta^3} \approx 0.41\rho_{\rm c}\ ,$$

where $\rho_{\rm c}$ is the Planck density (1.9), and the value (7.72) has been used for the Barbero–Immirzi parameter β. It has been shown that the quantum operator corresponding to the density remains bounded on the physical Hilbert space of these models, and that the upper bound of its spectrum coincides with $\rho_{\rm c}$. These results strongly indicate the avoidance of classical singularities in loop quantum cosmology. Whether this avoidance survives in more general models and whether there are robust singularity resolution theorems remains open, however (Ashtekar and Singh 2011).

Further reading: Bojowald (2011), Coule (2005), Halliwell (1991).

9
String theory

9.1 General introduction

The approaches discussed so far start from the assumption that the gravitational field can be quantized on its own. That this can really be done is, however, not clear. One could imagine that the problem of quantum gravity can only be solved within a unified quantum framework of all interactions. The main candidate up to now to achieve this goal has been superstring theory. Our interest here is mainly in the role of quantum-gravitational aspects. It is not our aim to give an introduction to the many physical and mathematical aspects of string theory. This has been done in a series of excellent textbooks; see in particular Green *et al.* (1987), Lüst and Theisen (1989), Polchinski (1998*a,b*), Kaku (1999), Becker *et al.* (2007), and Zwiebach (2009). Mohaupt (2003) gives a concise overview with particular emphasis on gravitational aspects. For more details, we refer the reader to these references.

String theory began as an attempt to explain the spectrum of hadrons. After the discovery of quantum chromodynamics and its successful predictions, it was abandoned as such. It was, however, realized that string theory could in principle implement a theory of quantum gravity (Scherk and Schwarz 1974; Yoneya 1974). The main reason is the appearance of a massless spin-2 particle in the spectrum of the string. As we learned in Chapter 2, such a particle necessarily leads to GR in the low-energy limit.

String theory transcends the level of local field theory because its fundamental objects are one-dimensional entities ('strings') instead of fields defined at space–time points. More recently, it has turned out that higher-dimensional objects ('branes') appear within string theory in a natural way and on an equal footing with strings (see below). We shall nevertheless continue to talk about 'string theory'.

What are the main features of string theory?

1. String theory necessarily contains gravity. The graviton appears as an excitation of closed strings. Open strings do not contain the graviton by themselves, but since they contain closed strings as virtual contributions, the appearance of the graviton is unavoidable there too.
2. String theory necessarily leads to gauge theories, since the corresponding gauge bosons are found in the string spectrum.
3. String theory seems to need supersymmetry (SUSY) for a consistent formulation. Fermions are therefore an essential ingredient.
4. All 'particles' arise from string excitations. Therefore, they are no longer fundamental and their masses should in principle be fixed by the string mass scale.

5. Higher space–time dimensions appear in a natural way, thus implementing the old idea of Kaluza and Klein (or some modern variant).
6. As emphasized above, string theory entails a unified quantum description of all interactions.
7. Since one can get chiral gauge couplings from string theory, the hope is raised that one can derive the Standard Model of elementary particles from it (although we are still very far from achieving this goal).

The free bosonic string has already been introduced in Section 3.2. Its fundamental dimensionful parameter is α' or the string length $l_{\mathrm{s}} = \sqrt{2\alpha'\hbar}$ derived from it. In view of the unification idea, one would expect that l_{s} is roughly of the order of the Planck length l_{P}. The starting point is the Polyakov action S_{P} (3.45). Making use of the three local symmetries (two diffeomorphisms and one Weyl transformation), we put this action into the 'gauge-fixed form' (3.52). In this form, the action still possesses an invariance with respect to conformal transformation, which form an infinite-dimensional group in two dimensions. It therefore gives rise to an infinite number of generators—the generators L_n of the Virasoro algebra (3.55).

Consider an *open string* with ends at $\sigma = 0$ and $\sigma = \pi$. Variation of the action (3.52) yields, after a partial integration, the expression

$$\delta S_{\mathrm{P}} = \frac{1}{2\pi\alpha'}\int_{\mathcal{M}} \mathrm{d}^2\sigma \; \left(\eta^{\alpha\beta}\partial_\alpha\partial_\beta X_\mu\right)\delta X^\mu \\ + \int_{\partial\mathcal{M}} \mathrm{d}\tau \; \left(X'_\mu \delta X^\mu\big]_{\sigma=\pi} - X'_\mu \delta X^\mu\big]_{\sigma=0}\right). \tag{9.1}$$

The classical theory demands that $\delta S_{\mathrm{P}} = 0$. For the surface term to vanish, one has the following options. One is to demand that $X'_\mu = 0$ for $\sigma = 0$ and $\sigma = \pi$ (a *Neumann* condition). This would guarantee that no momentum exits from the ends of the string. Alternatively, one can demand the *Dirichlet* condition, $X_\mu =$ constant for $\sigma = 0$ and $\sigma = \pi$. This condition comes into play automatically if the duality properties of the string are taken into account; see Section 9.2.3. The vanishing of the worldsheet integral in the variation of S_{P} leads to the wave equation

$$\left(\frac{\partial^2}{\partial\tau^2} - \frac{\partial^2}{\partial\sigma^2}\right) X^\mu(\sigma,\tau) = 0. \tag{9.2}$$

The solution of this equation for Neumann boundary conditions is

$$X^\mu(\sigma,\tau) = x^\mu + 2\alpha' p^\mu \tau + \mathrm{i}\sqrt{2\alpha'}\sum_{n\neq 0}\frac{\alpha_n^\mu}{n}\mathrm{e}^{-\mathrm{i}n\tau}\cos n\sigma. \tag{9.3}$$

Here, x^μ and p^μ denote the position and momentum, respectively, of the centre of mass. These would be the only degrees of freedom for a point particle. The quantities α_n^μ are the Fourier components (oscillator coordinates) and obey $\alpha_{-n}^\mu = (\alpha_n^\mu)^\dagger$ due to the reality of the X^μ. The solution (9.3) describes a standing wave.

In view of the path-integral formulation, it is often convenient to continue the worldsheet formally into the Euclidean regime, that is, to introduce worldsheet coordinates $\sigma^1 := \sigma$ and $\sigma^2 := \mathrm{i}\tau$. One can then use the complex coordinate

$$z = \mathrm{e}^{\sigma^2 - \mathrm{i}\sigma^1}, \tag{9.4}$$

with respect to which the solution (9.3) reads

$$X^\mu(z, \bar{z}) = x^\mu - \mathrm{i}\alpha' p^\mu \ln(z\bar{z}) + \mathrm{i}\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{\alpha_n^\mu}{n} \left(z^{-n} + \bar{z}^{-n}\right). \tag{9.5}$$

For a *closed string*, the boundary condition $X^\mu(\sigma) = X^\mu(\sigma + 2\pi)$ is sufficient. The solution of (9.2) can then be written as

$$X^\mu(\sigma, \tau) = X_\mathrm{R}^\mu(\sigma^-) + X_\mathrm{L}^\mu(\sigma^+), \tag{9.6}$$

where we have introduced the lightcone coordinates $\sigma^+ := \tau + \sigma$ and $\sigma^- := \tau - \sigma$. The indices R and L correspond to modes which would appear to be 'right-moving' and 'left-moving', respectively, in a two-dimensional space–time diagram. Explicitly, one has

$$X_\mathrm{R}^\mu(\sigma^-) = \frac{x^\mu}{2} + \frac{\alpha'}{2} p^\mu \sigma^- + \mathrm{i}\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{\alpha_n^\mu}{n} \mathrm{e}^{-\mathrm{i}n\sigma^-} \tag{9.7}$$

and

$$X_\mathrm{L}^\mu(\sigma^+) = \frac{x^\mu}{2} + \frac{\alpha'}{2} p^\mu \sigma^+ + \mathrm{i}\sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} \frac{\tilde{\alpha}_n^\mu}{n} \mathrm{e}^{-\mathrm{i}n\sigma^+}, \tag{9.8}$$

where $\tilde{\alpha}_n^\mu$ denotes the Fourier components of the left-moving modes. One can also give a formulation with respect to z and $\bar{z}$, but this will be omitted here. It is convenient to define

$$\tilde{\alpha}_0^\mu = \alpha_0^\mu = \sqrt{\frac{\alpha'}{2}} p^\mu$$

for the closed string, and

$$\alpha_0^\mu = \sqrt{2\alpha'} p^\mu$$

for the open string.

In Section 3.2, we introduced the string Hamiltonian; see (3.53). Inserting the classical solution for X^μ into this expression, we obtain

$$H = \frac{1}{2} \sum_{n=-\infty}^{\infty} \alpha_{-n}\alpha_n \tag{9.9}$$

for the open string, and

$$H = \frac{1}{2} \sum_{n=-\infty}^{\infty} \left(\alpha_{-n}\alpha_n + \tilde{\alpha}_{-n}\tilde{\alpha}_n\right) \tag{9.10}$$

for the closed string, where $\alpha_{-n}\alpha_n$ is a shorthand for $\eta_{\mu\nu}\alpha^\mu_{-n}\alpha^\nu_n$, etc. In (3.54), we introduced the quantities L_m for the open string; one has in particular $L_0 = H$. For the closed string, one can analogously define

$$L_m = \frac{1}{\pi\alpha'}\int_0^{2\pi} \mathrm{d}\sigma\ \mathrm{e}^{-\mathrm{i}m\sigma} T_{--} = \frac{1}{2}\sum_{n=-\infty}^{\infty} \alpha_{m-n}\alpha_n, \tag{9.11}$$

$$\tilde{L}_m = \frac{1}{\pi\alpha'}\int_0^{2\pi} \mathrm{d}\sigma\ \mathrm{e}^{\mathrm{i}m\sigma} T_{++} = \frac{1}{2}\sum_{n=-\infty}^{\infty} \tilde{\alpha}_{m-n}\tilde{\alpha}_n, \tag{9.12}$$

from which one obtains $H = L_0 + \tilde{L}_0$. As we have seen in Section 3.2, the L_m (and $\tilde{L}_m$) vanish as constraints.

Recalling that $\alpha_0^\mu = \sqrt{2\alpha'}p^\mu$ for the open string and, therefore,

$$\alpha_0^2 \equiv \alpha_0^\mu \alpha_0^\nu \eta_{\mu\nu} = 2\alpha' p^\mu p_\mu \equiv -2\alpha' M^2,$$

the condition $0 = L_0 = H$ enables us to obtain the following expression for the mass M of the open string as a function of the oscillatory string modes:

$$M^2 = \frac{1}{\alpha'}\sum_{n=1}^{\infty} \alpha_{-n}\alpha_n. \tag{9.13}$$

From $H = L_0 + \tilde{L}_0$, we obtain the following expression similarly for the mass of the closed string:

$$M^2 = \frac{2}{\alpha'}\sum_{n=1}^{\infty} \left(\alpha_{-n}\alpha_n + \tilde{\alpha}_{-n}\tilde{\alpha}_n\right). \tag{9.14}$$

The variables x^μ, p^μ, α^μ_n, and $\tilde{\alpha}^\mu_n$ obey Poisson-bracket relations that follow from the fundamental Poisson brackets between the X^μ and their canonical momenta P^μ (Section 3.2). Upon quantization, one obtains (setting $\hbar = 1$)

$$[x^\mu, p^\nu] = \mathrm{i}\eta^{\mu\nu}, \tag{9.15}$$

$$[\alpha^\mu_m, \alpha^\nu_n] = m\delta_{m,-n}\eta^{\mu\nu}, \tag{9.16}$$

$$[\tilde{\alpha}^\mu_m, \tilde{\alpha}^\nu_n] = m\delta_{m,-n}\eta^{\mu\nu}, \tag{9.17}$$

$$[\alpha^\mu_m, \tilde{\alpha}^\nu_n] = 0. \tag{9.18}$$

The Minkowski metric $\eta^{\mu\nu}$ appears because of Lorentz invariance. It can give rise to negative probabilities, which must be carefully avoided in the quantum theory.

The next task is to construct a Fock space out of the vacuum state $|0, p^\mu\rangle$, which is the ground state of a single string with momentum p^μ, *not* the no-string state. The above algebra of the oscillatory modes can be written, after rescaling, as the usual oscillator algebra of annihilation and creation operators a^μ_m and $a^{\mu\dagger}_m$,

$$\alpha^\mu_m = \sqrt{m}\, a^\mu_m, \quad \alpha^\mu_{-m} = \sqrt{m}\, a^{\mu\dagger}_m \ , m > 0.$$

One therefore has

$$\alpha_m^\mu|0, p^\mu\rangle = \tilde{\alpha}_m^\mu|0, p^\mu\rangle = 0,\ m > 0, \tag{9.19}$$

and

$$\alpha_0^\mu|0, p^\mu\rangle = \tilde{\alpha}_0^\mu|0, p^\mu\rangle = \sqrt{\frac{\alpha'}{2}}p^\mu|0, p^\mu\rangle. \tag{9.20}$$

The rest of the spectrum is generated by the creation operators α_m^μ and $\tilde{\alpha}_m^\mu$ for $m < 0$. In order to implement the conformal generators L_n in the quantum theory, one must address the issue of operator ordering. It has already been mentioned in Section 3.2 that this leads to the presence of a central term in the quantum algebra; see (3.56). Consequently, one cannot impose equations of the form $L_n|\psi\rangle = 0$ for all n. This is different from the spirit of the Wheeler–DeWitt equation, for which all constraints are implemented in this form. Instead, one can achieve this here only for $n > 0$. The demand for the absence of a Weyl anomaly on the worldsheet (see the next section) fixes the number D of the embedding space–time to be equal to 26. This is also called the 'critical dimension'. For the critical dimension, one then gets the following expressions for the mass spectrum, which differ from their classical counterparts (9.13) and (9.14) by constants (see e.g. Polchinski 1998*a*). For the open string, one has

$$M^2 = \frac{1}{\alpha'}\left(\sum_{n=1}^{\infty}\alpha_{-n}\alpha_n - 1\right) =: \frac{1}{\alpha'}(N-1), \tag{9.21}$$

where N denotes the level of excitation, while for the closed string one has

$$M^2 = \frac{4}{\alpha'}\left(\sum_{n=1}^{\infty}\alpha_{-n}\alpha_n - 1\right) = \frac{4}{\alpha'}\left(\sum_{n=1}^{\infty}\tilde{\alpha}_{-n}\tilde{\alpha}_n - 1\right). \tag{9.22}$$

The ground state ($N = 0$) for the open string thus has

$$M^2 = -\frac{1}{\alpha'} < 0. \tag{9.23}$$

(In D dimensions, one would have $M^2 = (2 - D)/24\alpha'$.) The corresponding particle describes a *tachyon*—a particle with negative mass-squared—which signals the presence of an unstable vacuum. There may be a different vacuum which is stable, and there is indeed some evidence that this is the case for the bosonic string; cf. Berkovits *et al.* (2000). What is clear is that the presence of SUSY eliminates the tachyon. This is one of the main motivations for introducing the *superstring* (Section 9.2.4). For the first excited state, one finds

$$|e, p^\nu\rangle = e_\mu\alpha_{-1}^\mu|0, p^\nu\rangle \tag{9.24}$$

with a polarization vector e_μ that turns out to be transverse to the string propagation, $e_\mu p^\mu = 0$, and therefore corresponds in the critical dimension to $D - 2 = 24$ degrees of freedom. Since $M^2 = 0$ for $N = 1$, this state describes a massless vector boson (a 'photon'). In fact, it turns out that had we chosen $D \neq 26$, we would have encountered a breakdown of Lorentz invariance. For $N > 1$, excited states correspond to massive particles. They are usually neglected because their masses are assumed to be of the order of the Planck mass—this is the mass scale of unification, where string theory is of relevance (since we expect a priori that $l_s \sim l_P$).

We emphasize that here we are dealing with higher-dimensional representations of the Poincaré group, which do not necessarily have analogues in $D = 4$. Therefore, the usual language of photons, etc., should not be taken literally.

For the closed string, there is the additional restriction $L_0 = \tilde{L}_0$, leading to

$$\sum_{n=1}^{\infty} \alpha_{-n}\alpha_n = \sum_{n=1}^{\infty} \tilde{\alpha}_{-n}\tilde{\alpha}_n .$$

The ground state is again a tachyon, with mass squared $M^2 = -4/\alpha'$. The first excited state is massless, $M^2 = 0$, and is described by

$$|e, p^\nu\rangle = e_{\mu\nu}\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1}|0, p^\nu\rangle, \tag{9.25}$$

where $e_{\mu\nu}$ is a transverse polarization tensor, $p^\mu e_{\mu\nu} = 0$. The state (9.25) can be decomposed into its irreducible parts. One thereby obtains a symmetric traceless tensor, a scalar, and an antisymmetric tensor. The symmetric tensor describes a spin-2 particle in $D = 4$ and can therefore—in view of the uniqueness features discussed in Chapter 2—be identified with the *graviton*. It is at this stage that string theory makes its first contact with quantum gravity. The perturbation theory discussed in Chapter 2 will thus be implemented in string theory. But, as we shall see in the next section, string theory can go beyond it.

The scalar is usually referred to as the *dilaton*, Φ. In $D = 4$, the antisymmetric tensor also has spin zero and in this case is called the *axion*; in the general case, it is called the Kalb–Ramond field. The fact that massless fields appear in the open- and the closed-string spectrum is very interesting. Both the massless vector boson and the graviton couple to conserved currents and thereby introduce the principle of gauge invariance into string theory. For closed strings, higher excited states also lead to massive ('heavy') particles.

Up to now we have discussed oriented strings, that is, strings whose quantum states have no invariance under $\sigma \to -\sigma$. We note that one can also have non-oriented strings by requiring this invariance to hold. For closed strings, this invariance would correspond to an exchange between right- and left-moving modes. It turns out that the graviton and the dilaton are also present for non-oriented strings, but not the axion.

9.2 Quantum-gravitational aspects

9.2.1 The Polyakov path integral

We have seen in the last section that the graviton appears in a natural way in the spectrum of closed strings. Linearized quantum gravity is, therefore, automatically contained in string theory. Here we discuss other aspects which are relevant in the context of quantum gravity.

One can generalize the Polyakov action (3.45) to the situation of a string moving in a general D-dimensional curved space–time. It makes sense to take into account not only gravity but also the other massless fields that arise in string excitations—the dilaton and the 'axion'. One therefore formulates the generalized Polyakov action as

$$\begin{aligned} S_{\mathrm{P}} &:= S_\sigma + S_\phi + S_B \\ &= -\frac{1}{4\pi\alpha'}\int \mathrm{d}^2\sigma \, \Big(\sqrt{h}h^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X^\nu g_{\mu\nu}(X) \\ &\qquad -\alpha'\sqrt{h}\,{}^{(2)}R\Phi(X) + \epsilon^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X^\nu B_{\mu\nu}(X)\Big) . \end{aligned} \tag{9.26}$$

The fields $g_{\mu\nu}$ (D-dimensional metric of the embedding space), Φ (dilaton), and $B_{\mu\nu}$ (antisymmetric tensor field or Kalb–Ramond field) are *background fields*; that is, they will not be integrated over in the path integral. The fields X^μ again define the embedding of the worldsheet into the D-dimensional space, which is also called the 'target space'. For historical reasons, an action in which the coefficients of the kinetic term depend on the fields themselves (here, $g_{\mu\nu}$ depends on X) is called a *non-linear sigma model*. This is why the first part on the right-hand side of (9.26) is abbreviated as S_σ. The second part S_ϕ is, in fact, independent of the string parameter α', since in natural units for which $\hbar = 1$, the dilaton is dimensionless. We emphasize that (9.26) describes a quantum field theory on the worldsheet, not in the target space. For the latter, one uses an effective action (see below).

In string theory, it has proved fruitful to employ a path-integral approach in which the worldsheet is taken to be Euclidean instead of Minkowskian; cf. Section 2.2. This has the advantage that the integral over the metric is better defined. Polchinski (1998*a*, p. 82) presents a formal argument why the resulting theory is equivalent to the original Minkowskian version.

In the Euclidean formulation, in which $\sigma^1 = \sigma$ and $\sigma^2 = \mathrm{i}\tau$, the starting point is thus

$$Z = \int \mathcal{D}X\mathcal{D}h \; \mathrm{e}^{-S_{\mathrm{P}}}, \tag{9.27}$$

where X and h are a shorthand for the embedding variables and the worldsheet metric, respectively. Only these variables are to be integrated over. In order to get a sensible expression, one must employ the gauge-fixing procedure outlined in Section 2.2.3. The invariances on the worldsheet involve two local diffeomorphisms and one Weyl transformation. Since $h_{ab}(\sigma^1, \sigma^2)$ has three independent parameters, one can *fix* it to a given 'fiducial' form $\tilde{h}_{ab}$, for example $\tilde{h}_{ab} = \delta_{ab}$ ('flat gauge') or $\tilde{h}_{ab} = \exp[2\omega(\sigma^1, \sigma^2)]\delta_{ab}$ ('conformal gauge'). As discussed in Section 2.2.3, the Faddeev–Popov determinant can be written as a path integral over (anticommuting) ghost fields. The action in (9.27) must then be replaced by the full action $S_{\mathrm{P}} + S_{\mathrm{ghost}} + S_{\mathrm{gf}}$, that is, augmented by ghost and gauge-fixing actions.

The full action is invariant under BRST transformations, which were briefly mentioned in Section 2.2.3. BRST invariance is an important concept, since it encodes the information about gauge invariance at the gauge-fixed level. For this reason, we shall give a brief introduction here (see e.g. Weinberg 1996 for more details). BRST transformations mix commuting and anticommuting fields (ghosts) and are generated by the 'BRST charge' Q_{B}. Let ϕ_a be a general set of first-class constraints (see Section 3.1.2),

$$\{\phi_a, \phi_b\} = f^c_{ab}\phi_c. \tag{9.28}$$

The BRST charge is then

$$Q_{\mathrm{B}} = \eta^a \phi_a - \tfrac{1}{2} P_c f^c_{ab} \eta^b \eta^a, \tag{9.29}$$

where η^a denotes the Faddeev–Popov ghosts and P_a their canonically conjugate momenta ('anti-ghosts') obeying $[\eta^a, P_b]_+ = \delta^a_b$. We have assumed here that the physical fields are bosonic; for a fermion, there would be a plus sign in (9.29). One can show that Q_{B} is nilpotent,

$$Q_{\mathrm{B}}^2 = 0. \tag{9.30}$$

This follows from (9.28) and the Jacobi identities for the structure constants.

In the quantum theory, BRST invariance of the path integral leads to the requirement that physical states should be BRST-invariant, that is,

$$\hat{Q}_{\mathrm{B}} |\Psi\rangle = 0. \tag{9.31}$$

This condition is less stringent than the Dirac condition, which states that physical states be annihilated by all constraints. Equation (9.31) can be satisfied for the quantized bosonic string, which is not the case for the Dirac conditions (see Section 3.2). The quantum version of (9.30) is

$$\left[\hat{Q}_{\mathrm{B}}, \hat{Q}_{\mathrm{B}}\right]_+ = 0. \tag{9.32}$$

For this to be satisfied, the total central charge of the X^μ-fields and the Faddeev–Popov ghosts must vanish. Since it turns out that the ghosts have central charge -26, this leads to the condition

$$c_{\mathrm{tot}} = c + c_{\mathrm{ghost}} = D - 26 = 0. \tag{9.33}$$

The string must therefore move in 26 dimensions! In the case of the superstring (see Section 9.2.4), the corresponding condition leads to $D = 10$. The condition (9.32) thus carries information about quantum anomalies (here, the Weyl anomaly) and the possibility of cancelling them by ghosts.[1] One can prove the 'no-ghost theorem' (see e.g. Polchinski 1998*a*): the Hilbert space arising from BRST quantization has a positive inner product and is isomorphic to the Hilbert space of transverse string excitations.

How can the path integral (9.27) be evaluated? The sum over all 'paths' contains, in particular, a sum over all worldsheets, that is, a sum over all Riemann surfaces. In this sum, all topologies have to be taken into account. As an example, Fig. 9.1 shows the first two topologies that arise in the scattering of two closed strings. It is in this way that string *interactions* arise—as amplitudes in the path integral. Unlike the situation in four dimensions, the classification of these surfaces in two dimensions is well known. Consider as an example the dilaton part of the action (9.26),

$$S_\phi = \frac{1}{4\pi} \int \mathrm{d}^2\sigma\, \sqrt{h}\, {}^{(2)}\!R\, \Phi(X). \tag{9.34}$$

[1] One can also discuss non-critical strings living in $D \neq 26$ dimensions. They have a Weyl anomaly, which means that different gauge choices are inequivalent.

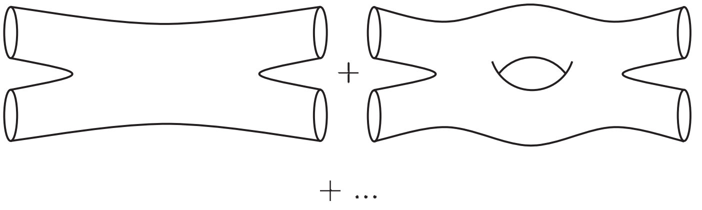

Fig. 9.1 The first two contributions to the scattering of two closed strings.

If Φ were constant, $\Phi(X) = \lambda$, this would yield

$$S_\phi = \chi\lambda = \lambda(2 - 2\mathfrak{g}), \tag{9.35}$$

where χ is the Euler number and $\mathfrak{g}$ the genus of the surface. (We assume for simplicity here that only handles are present, and no holes or cross-caps.) This then gives the contribution

$$\mathrm{e}^{-2\lambda(1-\mathfrak{g})} = \alpha^{\mathfrak{g}-1}$$

to the path integral, where we have introduced

$$\alpha := \mathrm{e}^{2\lambda} =: g_{\mathrm{c}}^2, \tag{9.36}$$

which plays the role of the 'fine-structure constant' for the loop expansion; g_{c} denotes the string-coupling constant for closed strings. Adding a handle corresponds to emission and reabsorption of a closed string.[2] The parameter g (meaning g_{c} or g_{o}, depending on the situation) is the expansion parameter for string loops. It must be emphasized that there is only *one* diagram at each order of the perturbation theory, in contrast to Feynman diagrams in quantum field theory. The reason is that point-like interactions are avoided. Such a 'smearing' can be done consistently in string theory, and it somewhat resembles the 'smearing' of the spin-network states discussed in Section 6.1. In this way, the usual divergences of quantum field theory seem to be avoided, although no proof is yet known that demonstrates finiteness at all loop orders. The sum as a whole does not converge and is not even Borel summable (i.e. the terms in the sum increase with $n!$, where n is the number of handles); see Gross and Periwal (1988). One therefore expects that, as in QED, this is an asymptotic series and thus must be an approximation to some non-perturbative theory.

In discussing scattering amplitudes, one must also specify the ingoing and outgoing string states, which must be given at infinity in the spirit of an 'S-matrix'. This is done with the help of 'vertex operators': in the example of Fig. 9.1, such an operator would correspond to four point-like insertions in the worldsheet. Vertex operators do not describe interactions but instead the creation or annihilation of a string state at a position on the worldsheet. Vertex operators must be included in the expression for the path integral.

It must also be emphasized that the gauge choice fixes the worldsheet metric only locally (e.g. to $\tilde{h}_{ab} = \delta_{ab}$). There may, however, be additional global degrees of freedom

[2]For the open string, one finds $g_{\mathrm{o}}^2 \propto \mathrm{e}^{\lambda}$.

described by a finite number of parameters. These parameters are called *moduli*. In the case of the torus ($\mathfrak{g} = 1$), for example, this is the Teichmüller parameter or modulus $\tau \in \mathbb{C}$. These parameters must be summed over in the path integral.

For a string propagating in flat space–time, the demand for the absence of a Weyl anomaly leads to the restriction $D = 26$. What about the string in a curved space–time as described by (9.26)? At the tree level, the requirement that no Weyl anomaly be present on the worldsheet leads to an additional set of consistency equations; they follow from the vanishing of renormalization-group beta functions,

$$0 = R_{\mu\nu} - \frac{1}{4} H_\mu{}^{\lambda\rho} H_{\nu\lambda\rho} + 2\nabla_\mu \nabla_\nu \Phi + \mathcal{O}(\alpha'), \tag{9.37}$$

$$0 = \nabla_\lambda H^\lambda{}_{\mu\nu} - 2\nabla^\lambda \Phi H_{\lambda\mu\nu} + \mathcal{O}(\alpha'), \tag{9.38}$$

$$0 = \frac{D-26}{6\alpha'} + \nabla_\mu \nabla^\mu \Phi - \frac{1}{2} \nabla_\mu \Phi \nabla^\mu \Phi - \frac{1}{24} H_{\mu\nu\rho} H^{\mu\nu\rho} + \mathcal{O}(\alpha'). \tag{9.39}$$

Here we have introduced the field strength $H_{\mu\nu\rho}$ associated with the antisymmetric tensor field,

$$H_{\mu\nu\rho} = 3!\ \partial_{[\mu} B_{\nu\rho]}.$$

The above set of equations (which in the highest order correspond to the Einstein equations in the presence of the dilaton and the antisymmetric tensor field) builds a bridge to the concept of effective action in string theory. As in Section 2.2.3, effective actions are useful for connecting the full theory with phenomenology.

9.2.2 Effective actions

The consistency equations (9.37)–(9.39) arise as field equations from the following *effective action* in D space–time dimensions:

$$S_{\text{eff}} = \frac{1}{2\kappa_0^2} \int \mathrm{d}^D x\ \sqrt{-g}\ \mathrm{e}^{-2\Phi} \Big(R - \frac{2(D-26)}{3\alpha'} - \frac{1}{12} H_{\mu\nu\rho} H^{\mu\nu\rho}$$
$$+ 4\nabla_\mu \Phi \nabla^\mu \Phi + \mathcal{O}(\alpha') \Big). \tag{9.40}$$

Higher orders include powers of α' and are thus genuine string corrections. They contain, for example, higher curvature terms, which means that in (9.37) one must make the replacement

$$R_{\mu\nu} \to R_{\mu\nu} + \frac{\alpha'}{2} R_{\mu\kappa\lambda\tau} R_\nu{}^{\kappa\lambda\tau} + \ldots.$$

The approximation of a classical space–time metric is well defined only if the curvature scale r_{c} associated with it obeys

$$r_{\mathrm{c}} \gg l_{\mathrm{s}}.$$

It has been speculated that a 'non-metric phase' will appear if this condition is violated (see e.g. Horowitz (1990), Greene (1997), and the references therein). The expansion of the effective action in powers of α' is thus a low-energy expansion.

One can compare the effective action with a Jordan–Brans–Dicke type of action, which contains an additional scalar field in the gravitational sector. There one has a kinetic term of the form $-4\omega\nabla_\mu\Phi\nabla^\mu\Phi$, where ω denotes the Brans–Dicke parameter (GR is recovered for $\omega \to \infty$). Comparison with (9.40) shows that string theory would correspond to $\omega = -1$. If the field Φ were really massless—as suggested by (9.40)—this would be in conflict with observations because the additional interaction of matter fields with Φ *in addition* to the metric would violate the equivalence principle (cf. Lämmerzahl 2003). The latter has been tested with great accuracy. One would, however, expect that the dilaton gets a mass term from the higher-order terms in α', so that no conflict with observation would arise. Because of the natural occurrence of the dilaton in string theory, theories with a scalar field in addition to the metric ('scalar–tensor theories') are widely studied; see Fuji and Maeda (2003) and Capozziello and Faraoni (2011).

We emphasize that (9.37) are the Einstein equations describing the coupling of the dilaton and the axion to the metric. It is interesting that these space–time equations follow from the Weyl invariance on the worldsheet. This gives one of the most important connections between the string and gravity.

One can also perform a Weyl transformation in space–time in order to put $S_{\rm eff}$ into a form in which the first term is just the space–time Ricci scalar without the dilaton. This is sometimes called the 'Einstein frame', in contrast to the 'string' or 'Jordan frame' of (9.40). While this new form is convenient for various situations, the physical form is given by (9.40), as can be seen from the behaviour of test particles—the physical metric is $g_{\mu\nu}$, not the Weyl-transformed metric. The Jordan frame is also distinguished by the fact that the standard fields are coupled minimally to the metric.

We have already emphasized above that the space–time metric, dilaton, and axion play only the role of background fields. The simplest solution for them is

$$g_{\mu\nu} = \eta_{\mu\nu}\ , \quad B_{\mu\nu} = 0\ , \quad \Phi = \text{const.} = \lambda.$$

It is usually claimed that, quite generally, the stationary points of $S_{\rm eff}$ correspond to possible ground states ('vacua') of the theory. String theory may, in fact, predict a huge number of such vacua; cf. Douglas (2003). The space of all string-theory vacua is also called the 'landscape' (Susskind 2003). A selection criterion for the most probable wave function propagating on such a landscape background is discussed in Mersini-Houghton (2005).

It is clear from (9.39) that $D = 26$ is a necessary condition for the solution with constant background fields. Thus, we have recovered the old consistency condition for the string in flat space–time. However, there are now solutions of (9.39) with $D \neq 26$ and $\Phi \neq$ constant, which would correspond to a solution with a large cosmological constant $\propto (D-26)/6\alpha'$, in conflict with observation.

The parameter κ_0 in (9.40) does not have a physical significance by itself, since it can be changed by a shift in the dilaton. The physical gravitational constant (in D dimensions) is

$$16\pi G_D = 2\kappa_0^2 \mathrm{e}^{2\lambda}. \tag{9.41}$$

Apart from α'-corrections, one can also consider loop corrections to (9.40). Since $g_{\rm c}$ is determined by the value of the dilaton (see (9.36)), the tree-level action (9.40) is of order g_c^{-2}. The one-loop approximation is obtained at order g_c^0, the two-loop approximation at order g_c^2, and so on.[3]

In Section 9.1, we saw that the graviton appears as an excitation mode for closed strings. What is the connection to the appearance of gravity in the effective action (9.40)? Such a connection is established through the ansatz

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \sqrt{32\pi G} f_{\mu\nu}$$

(cf. (2.77)) and by making a perturbation expansion in the effective action with respect to $f_{\mu\nu}$. It then turns out that the term of order $f_{\mu\nu}$ yields the vertex operator for the string graviton state (see e.g. Mohaupt 2003). Moreover, it is claimed that exponentiating this graviton vertex operator leads to a 'coherent state' of gravitons. The connection between the graviton as a string mode and gravity in the effective action thus proceeds via a comparison of scattering amplitudes. For example, the amplitude for graviton–graviton scattering from the scattering of strings at tree level coincides with the field-theoretic amplitude of the corresponding process at tree level when derived from S_{eff}. The reason for this coincidence is the vanishing of the Weyl anomaly for the worldsheet. The coincidence continues to hold at higher loop orders and at higher orders in $\alpha' \sim l_s^2$. At tree level, one can explicitly show that there is a coincidence of the graviton–graviton scattering cross-section in type II superstring theory with the corresponding process in quantum GR (Sannan 1986); differences arising from SUSY only come into play when loops are involved. The cross-section is thus given (for the helicities considered there) by the result (2.102).

Since the string amplitude contains the parameter α' and the effective action contains the gravitational constant G_D, the comparison of the amplitudes yields a connection between the two; see, for example, Veneziano (1993). This connection is

$$G_D \sim g^2 l_s^{D-2}, \tag{9.42}$$

where g is again the string-coupling constant (here we do not distinguish between open and closed strings, and write for simplicity just g for the string coupling). An analogous relation holds between gauge couplings for grand unified theories and the string length.

Since we do not live in 26 dimensions, a connection must be made to the four-dimensional world. This is usually done through compactification of the additional dimensions, which have the form of 'Calabi–Yau manifolds'; cf. Candelas *et al.* (1985). In this way, one obtains a relation between the four-dimensional gravitational constant and the string length,

$$G \sim g^2 l_s^2, \tag{9.43}$$

in which the details of the compactification enter into geometric factors. Ideally, one would like to recover in this way other parameters such as particle masses or the number of families from the details of the compactification. However, this goal is still very distant.

The finiteness of the string length l_s leads to an effective cut-off at high momenta. It thus seems impossible to resolve arbitrarily small distances in an operational sense. In

[3] For open strings, odd orders of the coupling (g_o) also appear.

fact, using scattering gedanken experiments, one can derive a generalized uncertainty relation of the form

$$\Delta x > \frac{\hbar}{\Delta p} + \frac{l_{\mathrm{s}}^2}{\hbar}\Delta p \; ; \tag{9.44}$$

cf. Veneziano (1993) and the references therein. This matches the idea of a minimal length, which we also encountered in the canonical approach (Section 6.2), although D-branes (Section 9.2.3), for example, can probe smaller scales.

The relation (9.44) can be phenomenologically extended to a general class of uncertainty relations, also called the generalized uncertainty principle (GUP). One can calculate from the GUP various corrections to observed phenomena such as the Lamb shift. Up to now, such corrections have turned out either to be too small or to lead to upper bounds on the parameters present in the GUP (Das and Vagenas 2008).

How many fundamental constants appear in string theory? This has been a matter of some debate; see Duff *et al.* (2002). We adopt here the standpoint already taken in Chapter 1 that three dimensionful constants are needed, which can be taken to be c, $\hbar$, and l_{s}.

9.2.3 T-duality and branes

In this subsection, we shall introduce the concept of T-duality, from which one is led in a natural way to the concept of D-branes. The 'T' arises from the fact that one assumes here that the higher dimensions are compactified on tori.

The classical solutions for the closed string are given in Section 9.1. In order to introduce the concept of duality, we take the left- and right-moving modes of the closed string as independent; that is, we write

$$X^\mu_{\mathrm{R}}(\sigma^-) = \frac{x^\mu}{2} + \sqrt{\frac{\alpha'}{2}}\alpha_0^\mu(\tau-\sigma) + \cdots \, , \tag{9.45}$$

$$X^\mu_{\mathrm{L}}(\sigma^-) = \frac{\tilde{x}^\mu}{2} + \sqrt{\frac{\alpha'}{2}}\tilde{\alpha}_0^\mu(\tau+\sigma) + \cdots \, , \tag{9.46}$$

where $\cdots$ stands for 'oscillators' (this part does not play a role in the following discussion). The sum of the two thus yields

$$X^\mu = \frac{1}{2}\left(x^\mu + \tilde{x}^\mu\right) + \sqrt{\frac{\alpha'}{2}}\left(\alpha_0^\mu + \tilde{\alpha}_0^\mu\right)\tau + \sqrt{\frac{\alpha'}{2}}\left(\tilde{\alpha}_0^\mu - \alpha_0^\mu\right)\sigma + \cdots \, . \tag{9.47}$$

The oscillators are invariant under $\sigma \to \sigma + 2\pi$, but the X^μ transform as

$$X^\mu \to X^\mu + 2\pi\sqrt{\frac{\alpha'}{2}}\left(\tilde{\alpha}_0^\mu - \alpha_0^\mu\right) . \tag{9.48}$$

We now distinguish between a non-compact direction of space and a compact direction; see in particular Polchinski (1998*a*) for the following discussion. In the non-compact directions, the X^μ must be unique. From (9.48), one then obtains for them

$$\tilde{\alpha}_0^\mu = \alpha_0^\mu = \sqrt{\frac{\alpha'}{2}}p^\mu . \tag{9.49}$$

This is the situation encountered before, and we are back to (9.7) and (9.8). For a compact direction, the situation is different. Assume that there is one compact direction with radius R in the direction $\mu = 25$. The coordinate $X^{25} := X$ thus has period $2\pi R$. Under $\sigma \to \sigma + 2\pi$, X can now change by $2\pi w R$, $w \in \mathbb{Z}$, where w is called the 'winding number'. These modes are called 'winding modes' because they can wind around the compact dimension. Since $\exp(2\pi \mathrm{i} R p^{25})$ generates a translation around the compact dimension that must lead to the same state, the momentum $p^{25} := p$ must be discretized,

$$p = \frac{n}{R}\,, \quad n \in \mathbb{Z}. \tag{9.50}$$

From

$$p = \frac{1}{\sqrt{2\alpha'}}(\tilde{\alpha}_0 + \alpha_0)$$

(α_0 is a shorthand for α_0^{25}, etc.), one gets for these 'momentum modes' the relation

$$\tilde{\alpha}_0 + \alpha_0 = \frac{2n}{R}\sqrt{\frac{\alpha'}{2}}. \tag{9.51}$$

For the winding modes, one has from (9.48)

$$2\pi\sqrt{\frac{\alpha'}{2}}(\tilde{\alpha}_0 - \alpha_0) = 2\pi w R$$

and therefore

$$\tilde{\alpha}_0 - \alpha_0 = wR\sqrt{\frac{2}{\alpha'}}. \tag{9.52}$$

This then yields

$$\alpha_0 = \left(\frac{n}{R} - \frac{wR}{\alpha'}\right)\sqrt{\frac{\alpha'}{2}} =: p_{\mathrm{R}}\sqrt{\frac{\alpha'}{2}}, \tag{9.53}$$

$$\tilde{\alpha}_0 = \left(\frac{n}{R} + \frac{wR}{\alpha'}\right)\sqrt{\frac{\alpha'}{2}} =: p_{\mathrm{L}}\sqrt{\frac{\alpha'}{2}}. \tag{9.54}$$

For the mass spectrum, one obtains

$$\begin{aligned} M^2 = -\sum_{\mu\neq 25} p^\mu p_\mu + p^2 &= \frac{4}{\alpha'}\left(\sum_{n=1}^{\infty}\alpha_{-n}\alpha_n - 1\right) + \frac{2(\alpha_0)^2}{\alpha'} \\ &= \frac{4}{\alpha'}\left(\sum_{n=1}^{\infty}\tilde{\alpha}_{-n}\tilde{\alpha}_n - 1\right) + \frac{2(\alpha_0)^2}{\alpha'}. \end{aligned} \tag{9.55}$$

The expressions in parentheses correspond to the excitation levels of the right- and left-moving modes, respectively, in the non-compact dimensions.

Of particular interest are the limiting cases $R \to \infty$ and $R \to 0$. For $R \to \infty$, one has

$$\alpha_0 \to -\frac{wR}{\sqrt{2\alpha'}} \ , \quad \tilde{\alpha}_0 \to \frac{wR}{\sqrt{2\alpha'}}. \tag{9.56}$$

Since $M^2 \to \infty$ for $w \neq 0$, all states become infinitely massive in this case. For $w = 0$, in contrast, one gets a *continuum* of states for all n. For $R \to 0$, one gets

$$\alpha_0 \to \frac{n}{R}\sqrt{\frac{\alpha'}{2}} \ , \quad \tilde{\alpha}_0 \to \frac{n}{R}\sqrt{\frac{\alpha'}{2}}. \tag{9.57}$$

For $n \neq 0$, all states become infinitely massive, while for $n = 0$ there is a continuum for all w. These are the winding states that can wind around the extra dimension without any cost in terms of energy (note that $H_0 = (\alpha_0^2 + \tilde{\alpha}_0^2) \to 0$ for $n = 0$). This is a typical feature of string theory related to the presence of one-dimensional objects and has no counterpart in field theory. We recognize from (9.53) and (9.54) that there is a symmetry between R and the 'dual radius' $R_{\rm D}$,

$$R \leftrightarrow R_{\rm D} := \frac{\alpha'}{R} = \frac{l_{\rm s}^2}{2R}, \tag{9.58}$$

which corresponds to

$$n \leftrightarrow w \ , \quad \alpha_0 \leftrightarrow -\alpha_0 \ , \quad \tilde{\alpha}_0 \leftrightarrow \tilde{\alpha}_0. \tag{9.59}$$

The mass spectrum is identical in the two cases. The above duality between R and $R_{\rm D}$ is called 'T-duality'. It is an exact symmetry of perturbation theory (and beyond) for closed strings. The critical value is, of course, obtained for $R \sim R_{\rm D} \sim l_{\rm s}$, which is just the length of the string, as expected.

It is useful for the discussion below to go again to Euclidean space,

$$z = {\rm e}^{\sigma^2 - {\rm i}\sigma^1} \ , \quad \sigma^1 = \sigma, \ \sigma^2 = {\rm i}\tau.$$

Equation (9.47) then reads

$$\begin{aligned} X^\mu(z,\bar{z}) &= X_{\rm R}^\mu(z) + X_{\rm L}^\mu(\bar{z}) \\ &= \frac{x^\mu + \tilde{x}^\mu}{2} - {\rm i}\sqrt{\frac{\alpha'}{2}}(\tilde{\alpha}_0^\mu + \alpha_0^\mu)\sigma^2 + \sqrt{\frac{\alpha'}{2}}(\tilde{\alpha}_0^\mu - \alpha_0^\mu)\sigma^1 + \dots \\ &\xrightarrow{\rm T} -X_{\rm R}^\mu(z) + X_{\rm L}^\mu(\bar{z}). \end{aligned} \tag{9.60}$$

What happens for the open string? In the limit $R \to 0$, it has no possibility of winding around the compactified dimension and therefore seems to live only in $D-1$ dimensions. There are, however, closed strings present in the theory of open strings, and the open string can in particular have vibrations into the 25th dimension. Only the *endpoints* of the open string are constrained to lie on a $(D-1)$-dimensional hypersurface.

For the vibrational part in the 25th dimension, one can use the expression (9.60) for the closed string. The choice $X^{25} := X$ in this equation yields

$$X = \frac{x + \tilde{x}}{2} + \sqrt{\frac{\alpha'}{2}}(\tilde{\alpha}_0 - \alpha_0)\sigma^1 - \mathrm{i}\sqrt{\frac{\alpha'}{2}}(\tilde{\alpha}_0 + \alpha_0)\sigma^2 + \ldots, \tag{9.61}$$

so that its dual is

$$X_{\mathrm{D}} = \frac{x + \tilde{x}}{2} + \sqrt{\frac{\alpha'}{2}}(\tilde{\alpha}_0 + \alpha_0)\sigma^1 - \mathrm{i}\sqrt{\frac{\alpha'}{2}}(\tilde{\alpha}_0 - \alpha_0)\sigma^2 + \ldots, \tag{9.62}$$

that is, $X \leftrightarrow X_{\mathrm{D}}$ corresponds to $-\mathrm{i}\sigma^2 \leftrightarrow \sigma^1$, leading in particular to

$$X'_{\mathrm{D}} \equiv \frac{\partial X_{\mathrm{D}}}{\partial \sigma^1} \leftrightarrow \mathrm{i}\frac{\partial X}{\partial \sigma^2} \equiv \mathrm{i}\dot{X}.$$

Integration yields

$$X_{\mathrm{D}}(\pi) - X_{\mathrm{D}}(0) = \int_0^\pi \mathrm{d}\sigma^1\ X'_{\mathrm{D}} = \mathrm{i}\int_0^\pi \mathrm{d}\sigma^1\ \dot{X}.$$

Insertion of $\dot{X}$ from (9.3) yields

$$X_{\mathrm{D}}(\pi) - X_{\mathrm{D}}(0) = 2\pi\alpha' p = \frac{2\pi\alpha' n}{R} = 2\pi n R_{\mathrm{D}}\ ,\ n \in \mathbb{Z},$$

where only vibrational modes have been considered. The dual coordinates thus obey a Dirichlet-type condition (an exact Dirichlet condition would arise for $n = 0$; cf. the remarks before (9.2)). Since, therefore, $X_{\mathrm{D}}(\pi)$ and $X_{\mathrm{D}}(0)$ differ by a multiple of the internal circumference $2\pi R_{\mathrm{D}}$ of the dual space, the endpoints must lie on a $(D-1)$-dimensional hypersurface. Consideration of several open strings reveals that it is actually the same hypersurface. This hypersurface is called a D-*brane*, where 'D' refers to the Dirichlet condition that holds normal to the brane. Sometimes one refers to it more precisely as the 'Dp-brane', where p is the number of space dimensions.

A D-brane is a *dynamical* object, since momentum can leak out of the string and be absorbed by the brane (this cannot happen with a Neumann boundary condition, which still holds in the directions tangential to the plane). A D-brane is a soliton of string theory and can be described by an action that resembles an action proposed long ago by Born and Infeld to describe non-linear electrodynamics (it was then meant as a candidate for a modification of linear electrodynamics at short distances). A D-brane can carry generalized electric and magnetic charges.

The above discussion can be extended to the presence of gauge fields. This is because open strings allow additional degrees of freedom called 'Chan–Paton factors'. These are 'charges' i and j $(i, j = 1, \ldots, n)$ that reside at the endpoints of the string (historically, people were thinking about quark–antiquark pairs). One can introduce a $\mathrm{U}(n)$ symmetry acting on (and only on) these charges and arrive at the concept of $\mathrm{U}(n)$ gauge bosons living *on* the branes (n branes can exist at different positions).

Interestingly, n coincident D-branes give rise to '$n \times n$' matrices for the embedding variables X^μ and the gauge fields A_a. One thus arrives at space–time coordinates that do not commute, giving rise to the notion of non-commutative space–time. It has been argued that the D-brane action corresponds to a Yang–Mills action on a non-commutative worldvolume. Details are reviewed, for example, in Douglas and Nekrasov (2002).

The concept of D-branes is especially interesting in connection with its gravitational aspects. First, it plays a crucial role in the derivation of the black-hole entropy based on the counting of microscopic degrees of freedom (Section 9.2.5). Second, these branes allow one to localize gauge and matter fields on the branes, whereas the gravitational field can propagate through the full space–time. This gives rise to a number of interesting features discussed in the context of 'brane worlds'; cf. Section 9.2.6.

9.2.4 Superstrings

So far, we have not yet included fermions, which are necessary for a realistic description of the world. Fermions are implemented by the introduction of SUSY, which we have already discussed in Sections 2.3 and 5.3.6. In contrast to the discussion there, we shall introduce SUSY on the worldsheet here, not on space–time. This will help us to get rid of some problems of the bosonic string, such as the presence of tachyons. A string with SUSY is called a 'superstring'. Worldsheet SUSY will be related to space–time SUSY only indirectly here. We shall be brief in our discussion, and refer the reader to, for example, Polchinski (1998*b*) for more details.

We start by introducing the superpartners for the X^μ on the worldsheet—these are called ψ^μ_A, where μ is a space–time vector index and A is a worldsheet spinor index ($A = 1, 2$). The ψ^μ_A are taken to be Majorana spinors and thus have two real components,

$$\psi^\mu = \begin{pmatrix} \psi_- \\ \psi_+ \end{pmatrix}. \tag{9.63}$$

They are not to be confused with the gravitinos, which are trivial in two dimensions.

The 'superversion' of the (flat) Polyakov action (3.52) is the 'RNS action' (named after Ramond, Neveu, and Schwarz). It reads

$$S_{\text{RNS}} = -\frac{1}{4\pi\alpha'} \int_{\mathcal{M}} \mathrm{d}^2\sigma \ \left(\partial_\alpha X^\mu \partial^\alpha X_\mu + \mathrm{i}\bar{\psi}^\mu \rho^\alpha \partial_\alpha \psi_\mu\right), \tag{9.64}$$

where the ρ^α denote two-dimensional Dirac matrices (suppressing worldsheet spinor indices). Thus, they obey

$$\left[\rho^\alpha, \rho^\beta\right]_+ = 2\eta^{\alpha\beta}. \tag{9.65}$$

The RNS action is invariant under global SUSY transformations on the worldsheet. The classical equations of motion are (9.2) and

$$\rho^\alpha \partial_\alpha \psi^\mu = 0. \tag{9.66}$$

To make SUSY on the worldsheet explicit, one can formulate the theory in a two-dimensional 'superspace' described by worldsheet coordinates (σ, τ) and two additional Grassmann coordinates θ^A. This must not be confused with the superspace of canonical gravity discussed in Chapters 4 and 5! We shall not elaborate on the SUSY formalism here.

Compared with the bosonic string, new features arise in the formulation of the *boundary conditions*. Considering first the open string, the requirement of a vanishing surface term in the variation of the action allows two possible boundary conditions for the fermions:

1. *Ramond* (R) boundary conditions: ψ is periodic (on a formally doubled world-sheet), and the sum in the mode expansion is over $n \in \mathbb{Z}$.
2. *Neveu–Schwarz* (NS) boundary conditions: ψ is antiperiodic, and the sum is over $n \in \mathbb{Z} + 1/2$ (this is possible because a relative sign is allowed for spinors).

It turns out that for consistency one must really have both types of boundary condition in the theory, the R-sector and the NS-sector of Hilbert space.

For the closed string, one can require that ψ_+ and ψ_- are either periodic (R) or antiperiodic (NS). Since they are independent, one gets four types of boundary condition: R–R, NS–R, R–NS, and NS–NS. Again, these must all be taken into account for consistency.

What about quantization? From the X^μ part, one gets the commutation relations (9.16)–(9.18) for the α_n^μ and the $\tilde{\alpha}_n^\mu$ as before. For the superstring, one has in addition anticommutators for the fermionic modes. They are of the type

$$[b_m^\mu, b_n^\nu]_+ = \eta^{\mu\nu}\delta_{m,-n}, \tag{9.67}$$

etc. One finds a SUSY extension of the Virasoro algebra, which again has a central charge. The requirement of a vanishing Weyl anomaly leads this time to $D = 10$ dimensions for the superstring.

Consider first the case of closed strings.[4] The NS–NS sector yields, as in the bosonic case, a tachyonic ground state. At $M = 0$, one again finds a graviton, a dilaton, and an antisymmetric tensor field ('axion' in four dimensions). The R–R sector yields antisymmetric tensor gauge fields, while the NS–R sector gives space–time fermions. The R–NS sector contains an exchange of left- and right-moving fermions compared with the NS–R sector. Among these fermions are massless gravitinos. In this sense, string theory contains space–time supergravity (SUGRA); see Section 2.3. An important notion (for both open and closed strings) is the 'GSO projection' (named after Gliozzi, Scherk, and Olive). This removes the tachyon and makes the spectrum supersymmetric. Moreover, it must necessarily be implemented in the quantum theory. In the R–R sector, the GSO projection applied to ground states can yield states of the opposite or of the same chirality. In the first case, one talks about a type IIA superstring (which is non-chiral), and in the latter case, about a type IIB superstring (which is chiral). Types IIA and IIB are oriented closed superstrings with $N = 2$ SUSY. After the GSO projection, there is no longer a tachyon in the NS–NS sector, but one still has the graviton, the dilaton, and the axion as massless states (for both type IIA and type IIB). In the NS–R and the R–NS sectors, one is left with two gravitinos and two dilatinos (the SUSY partners of the dilaton), which have opposite chiralities for type IIA and the same chirality for type IIB.

In the case of open strings, one gets in the NS sector a tachyonic ground state and a massless gauge boson. In the R sector, all states are space–time spinors. Again,

[4]We neglect all massive states in our discussion.

one gets rid of the tachyon by applying the GSO projection. This leads to the type I superstring—the only consistent theory with open (and closed) strings (the strings here are non-oriented). It must have the gauge group SO(32) and has $N = 1$ SUSY. In the closed-string sector of the type I theory, one must project type IIB onto states which are invariant under worldsheet parity in order to get non-oriented strings. There remain the graviton, the dilaton, a two-form field, one gravitino, and one dilatino. From the open-string sector, one gets massless vector and spinor fields.

In addition to types I, IIA, and IIB, there exists a consistent hybrid construction for closed strings that combine the bosonic string with type II superstrings. This is referred to as the 'heterotic string'; the right-moving part is taken from type II and the left-moving part is from the bosonic string. It possesses $N = 1$ SUSY. There exist two different versions, corresponding to the gauge groups SO(32) and $E_8 \times E_8$. Anomaly-free chiral models for particle physics can thus be constructed from string theory for these gauge groups. This has raised the hope that the Standard Model of strong and electroweak interactions can be derived from string theory—a hope, however, which up to now has not been realized.

To summarize, (the weak-field limits of) *five* consistent string theories in $D = 10$ dimensions have been found. To find a theory in four dimensions, one has to invoke a compactification procedure. Since no principle has yet been found to fix this, there are plenty of consistent string theories in four dimensions and it is not clear which one to choose.

This is not, however, the end of the story. Type IIA theory also contains D0-branes ('particles'); cf. Witten (1995). If one has n such D0-branes, their mass M is given by

$$M = \frac{n}{g\sqrt{\alpha'}}. \tag{9.68}$$

In the perturbative regime $g \ll 1$, this state is very heavy, while in the strong-coupling regime $g \to \infty$ it becomes lighter than any perturbative excitation. The mass spectrum (9.68) resembles a Kaluza–Klein spectrum; cf. the beginning of Section 9.2.6. It thus signals the presence of an *11th dimension* with radius

$$R_{11} = g\sqrt{\alpha'}. \tag{9.69}$$

The 11th dimension cannot be seen in string perturbation theory, which is a perturbation theory for small g. Since $D = 11$ is the maximum dimension in which SUSY can exist, this suggests a connection with 11-dimensional SUGRA. It is generally believed that the five string theories are the perturbative limits of one fundamental theory, of which 11-dimensional SUGRA is a low-energy limit. This fundamental theory, about which little is known, is called *M-theory*; cf. Nicolai (1999). A particular proposal for this theory is 'matrix theory', which employs only a finite number of degrees of freedom connected with a system of D0-branes; see Steinacker (2010) for a review.[5] The $N \to \infty$ limit of an SU(N) matrix model has been proposed as a model for M-theory (Banks *et al.* 1997). This model had already been discussed earlier in connection

[5]The group-field theory approach mentioned at the end of Chapter 6 is a higher-dimensional generalization of matrix models.

with 11-dimensional supermembranes (de Wit *et al.* 1988). The fundamental scale of M-theory is the 11-dimensional Planck length.

Our understanding of M-theory is indeed very limited. It is, for example, not yet possible to give a full non-perturbative calculation of graviton–graviton scattering, one of the important processes in quantum gravity (see Chapter 2). It is not even clear that M-theory exists.

In Section 9.2.3, we discussed the notion of T-duality, which connects descriptions of small and large radii. There is a second important notion of duality, called 'S-duality', which relates the five consistent superstring theories to each other. Thereby, the weak-coupling sector ($g \ll 1$) of one theory can be connected to the strong-coupling sector ($g \gg 1$) of another (or the same) theory.

9.2.5 Black-hole entropy

In Section 7.3, we reviewed attempts to calculate the Bekenstein–Hawking entropy (7.23) by counting microscopic degrees of freedom in canonical quantum gravity; see also the remarks in Section 8.1.3 on the situation in (2+1)-dimensional gravity. What can string theory say about this issue? It turns out that one can give a microscopic account of *extremal* black holes and black holes that are close to extremality. 'Extremal' is meant here with respect to the generalized electric and magnetic charges that can be present in the spectrum of string theory. This is analogous to the situation for an extremal Reissner–Nordström black hole (Section 7.1). We shall be brief in the following, and refer the reader to, for example, Horowitz (1998), Peet (1998), and de Wit (2006) for reviews.

The key idea in the string calculation of (7.23) is the notion of S-duality discussed at the end of the last subsection. A central role is played by so-called 'BPS states' (named after Bogomolnyi, Prasad, and Sommerfield), which have the important property that they are invariant under a non-trivial subalgebra of the full SUSY algebra. As a consequence, their mass is fixed in terms of their charges, and their spectrum is preserved on the transition from the weak- to the strong-coupling limit of string theory. In the weak-coupling limit, a BPS state can describe a bound state of D-branes, whose entropy $S_{\rm s}$ can be easily calculated. In the strong-coupling limit, the state can describe an extremal black hole, whose entropy can be calculated by use of (7.23). Interestingly, the two calculations lead to the same result. This was first shown by Strominger and Vafa (1996) for an extremal hole in five dimensions. It has to be emphasized that all these calculations are made in the semiclassical regime, in which the black hole is not too small. The final evaporation has therefore not yet been addressed.

We give here only some heuristic arguments why this result can hold, and refer to the above references for details. The level density d_N of a highly excited string state with excitation level N is (for open strings) given in the limit $N \to \infty$ by

$$d_N \sim \mathrm{e}^{4\pi\sqrt{N}} \approx \mathrm{e}^{M/M_0}, \tag{9.70}$$

where (9.21) has been used, and

$$M_0 := \frac{1}{4\pi\sqrt{\alpha'}}. \tag{9.71}$$

The temperature $T_0 := M_0/k_{\mathrm{B}}$ connected with M_0 is called the 'Hagedorn temperature' or the 'temperature of hell' because the free energy diverges as T_0 is approached (signalling a phase transition).

The expression for d_N can be understood as follows. If one divides a string with energy M into two parts with energies M_1 and M_2, one would expect that $M = M_1 + M_2$. The number of states would then obey

$$d_N(M) = d_N(M_1)d_N(M_2) = d_N(M_1 + M_2),$$

from which a relation of the form (9.70) follows. Using the statistical-physics formula $d_N = \exp(S_{\mathrm{s}})$, one finds for the 'string entropy'

$$S_{\mathrm{s}} \propto M \propto \sqrt{N}. \tag{9.72}$$

Since the gravitational constant depends on the string coupling (see (9.43)), the effective Schwarzschild radius $R_{\mathrm{S}} = 2GM$ increases if g is increased, and a black hole can form if g becomes large enough. On the other hand, if one starts from a black hole and decreases g, one finds that a highly excited string state is formed once R_{S} is smaller than l_{s}; cf. Horowitz and Polchinski (1997). It seems at first that (9.72) contradicts the Bekenstein–Hawking entropy, which (for the Schwarzschild case) is proportional to M^2, not M. That this is not a problem follows from the g-dependence of the gravitational constant. Following Horowitz (1998), we compare the mass $M \sim \sqrt{N}/l_{\mathrm{s}}$ of a string with the mass $M_{\mathrm{BH}} \sim R_{\mathrm{S}}/G$ of the black hole when $R_{\mathrm{S}} \approx l_{\mathrm{s}}$,

$$M_{\mathrm{BH}} \sim \frac{R_{\mathrm{S}}}{G} \approx \frac{l_{\mathrm{s}}}{G} \sim M \sim \frac{\sqrt{N}}{l_{\mathrm{s}}},$$

which leads to $l_{\mathrm{s}}^2/G \sim \sqrt{N}$. The entropy of the black hole is then given by

$$S_{\mathrm{BH}} \sim \frac{R_{\mathrm{S}}^2}{G} \approx \frac{l_{\mathrm{s}}^2}{G} \sim \sqrt{N}$$

and is thus comparable to the string entropy (9.72). Strings therefore possess enough states to yield the Bekenstein–Hawking entropy. It is most remarkable that an exact calculation yields (for BPS states) an exact coincidence between the two entropies. The fact that $R_{\mathrm{S}} \approx l_{\mathrm{s}}$ in the above estimate does not mean that the black hole is small: by eliminating l_{s} in the above expressions in favour of the Planck length $l_{\mathrm{P}} \sim \sqrt{G}$, one finds

$$R_{\mathrm{S}} \sim N^{1/4} l_{\mathrm{P}}.$$

Since $N \gg 1$, the Schwarzschild radius R_{S} is much greater than the Planck length—again a consequence of the fact that G varies with g, while l_{s} is fixed.

The exact calculations mentioned above refer to extremal black holes, for which the Hawking temperature is zero (see Section 7.1). It has, however, been possible to generalize the result to near-BPS states. Hawking radiation is then non-vanishing and corresponds to the emission of a closed string from a D-brane. If the D-brane state is traced out, the radiation is described by a thermal state. This is in accordance with Section 7.2, where it was argued that unitarity is preserved for the full system and that the mixed appearance of Hawking radiation arises from quantum entanglement with an 'environment', leading to decoherence (cf. Section 10.1). In the present case, the environment would be a system of D-branes. One would thus not expect any information-loss paradox to be present. In fact, in their range of validity, the calculations in string theory preserve unitarity. This is also supported by the AdS/CFT correspondence (Section 9.2.8), which relates gravity to *unitary* conformal field theories.

It is interesting that these calculations can yield exact cross-sections for the black hole, including the greybody factor $\Gamma_{\omega l}$ in (7.14). Unfortunately, it is not yet possible to extend these exact results to generic black holes. An exact treatment of the Schwarzschild black hole, for example, remains elusive.

Sub-leading corrections to both $S_{\rm BH}$ and $S_{\rm s}$ have been calculated and shown to be connected with higher curvature terms; cf. Mohaupt (2001) and the references therein.

9.2.6 Brane worlds

String theory employs higher dimensions for its formulation. The idea of using higher dimensions for unified theories goes back to the pioneering work of Kaluza and Klein in the 1920s; see, for example, Lee (1984) for a collection of reviews and an English translation of the original papers. In the simplest version, there is one additional space dimension, which is compactified to a circle with circumference $2\pi R$. We label the usual four dimensions by coordinates x^μ and the fifth dimension by y. In such a scenario, one can easily get particle masses in four dimensions from a massless five-dimensional field. Assuming for simplicity that the metric is flat, the dynamical equation for a massless scalar field Φ in five dimensions is given by the wave equation

$$\Box_5 \Phi(x^\mu, y) = 0, \tag{9.73}$$

where $\Box_5$ is the five-dimensional d'Alembert operator. Performing a Fourier expansion with respect to the fifth dimension,

$$\Phi(x^\mu, y) = \sum_n \varphi_n(x^\mu) \mathrm{e}^{\mathrm{i}ny/R}\ ,\ n \in \mathbb{Z}, \tag{9.74}$$

one obtains for the $\varphi_n(x^\mu)$ an effective equation of the form

$$\left(\Box_4 - \frac{n^2}{R^2}\right) \varphi_n(x^\mu) = 0, \tag{9.75}$$

where $\Box_4$ is the four-dimensional d'Alembert operator. Equation (9.75) is nothing but the four-dimensional Klein–Gordon equation for a massive scalar field $\varphi_n(x^\mu)$ with mass

$$m_n = \frac{|n|}{R}. \tag{9.76}$$

From the four-dimensional point of view, one thus has a whole 'Kaluza–Klein tower' of particles with increasing masses. At low energies $E \lesssim 1/R$, the massive Kaluza–Klein modes remain unexcited and only the massless mode with $n = 0$ remains. The higher dimensions only show up above energies $1/R$. Since no evidence has yet been seen at accelerators for the massive modes, the fifth dimension must be very small, definitely smaller than about 10^{-17} cm.

The Kaluza–Klein scenario has been generalized in various directions. In Section 9.2.3, we have seen that the notion of T-duality in string theory gives rise to the concept of D-branes. Gauge and matter fields are localized on the brane, whereas gravity can propagate freely through the higher dimensions (the bulk). One can therefore assume that our observed four-dimensional world is, in fact, such a brane embedded in higher dimensions. This gives rise to various 'brane-world scenarios', which often are very loosely related to string theory itself, taking from it only the idea of a brane, without necessarily giving a dynamical justification from string theory. General reviews have been given by Rubakov (2001) and Maartens and Koyama (2010).

In one scenario, the so-called 'ADD' approach, named after the authors of Arkani-Hamed *et al.* (1998), the brane tension (the energy per unit three-volume of the brane) and therefore its gravitational field are neglected; see Arkani-Hamed *et al.* (1998) and Antoniadis *et al.* (1998). A key ingredient in this approach is to take the extra dimensions to be compact (as in the standard Kaluza–Klein approach) but *not* microscopically small. The justification for this is that only gravity can probe the extra dimensions, and the gravitational attraction has only been tested down to distances of about 0.2 mm. Any value for R with $R \lesssim 0.1$ mm is thus allowed.

One of the main motivations for coming up with such a scenario is the *hierarchy problem*. This is the problem of why the scale of the electroweak interaction (about 1 TeV) is so much smaller than the Planck scale $m_{\rm P} \approx 10^{19}$ GeV. Comparing the (reduced) Planck mass $M_{\rm P}$ (1.7) with the estimated mass of the Higgs field, $m_{\rm H} \sim$ 150 GeV, one has

$$\left(\frac{M_{\rm P}}{m_{\rm H}}\right)^2 \sim 10^{32}.$$

Why is this a problem? The Higgs fields couples to all massive fields with a strength proportional to this mass. One would thus expect large quantum corrections to the Higgs mass from the virtual contribution of the massive quantum-gravity particles to the loop diagrams of the Standard Model. Without further input, these large corrections can only be avoided by a careful fine-tuning between the corrections and the bare Higgs mass, which is mysterious. One possible solution to the hierarchy problem invokes supersymmetry; see, for example, Martin (1997). Another possibility is to look for a cosmological origin (Hogan 2000).

The models invoking extra dimensions approach the hierarchy problem from a different perspective. In the ADD model, the fields of the Standard Model are constrained to the brane, while gravity can also propagate through the bulk. It could thus happen that the Planck mass m_D for the D-dimensional higher-dimensional space–time is of the same order as the electroweak scale (about 1 TeV), and that gravity appears weak on our brane only because of its dilution over the bulk (gravitons can be emitted not only inside the brane but also into the bulk).

The scenario can be described as follows. One starts with the Einstein–Hilbert action in $D = 4 + n$ dimensions,

$$S_{\rm EH} = \frac{1}{16\pi G_D} \int {\rm d}^4x\, {\rm d}^n y\, \sqrt{-g_D}\, {}^{(D)}R, \tag{9.77}$$

where the index D refers to the corresponding quantities in D dimensions, and n is the number of extra dimensions. In natural units, the D-dimensional gravitational constant G_D has the physical dimensions of an inverse mass to the $n+2$th power. We can thus write[6]

$$G_D = \frac{1}{m_D^{D-2}} = \frac{1}{m_D^{n+2}},$$

but should note that additional numerical factors are sometimes introduced into this relation.

Assuming that the D-dimensional metric is (approximately) independent of the extra dimensions, labelled by y, one gets from (9.77) an effectively four-dimensional action,

$$S_{\rm EH} = \frac{V_n}{16\pi G_D} \int \mathrm{d}^4x\sqrt{-g_4}\,{}^{(4)}R, \tag{9.78}$$

where $V_n \sim R^n$ denotes the volume of the extra dimensions.[7] Comparison with the four-dimensional Einstein–Hilbert action (1.1) gives the connection between m_D and the four-dimensional Planck mass,

$$m_{\rm P}^2 = V_n m_D^{2+n}. \tag{9.79}$$

The four-dimensional Planck mass is thus large (compared with the weak scale) because the extra dimensions are large. The hierarchy problem is thereby transferred to a different problem: why is R so large? An explanation should come from a fundamental theory such as string theory.

Assuming torus compactification and that m_D is in the TeV range, one can find from (9.79) the corresponding size of the extra dimensions. For $n = 1$, one finds that R is the size of the Solar System, which is definitely excluded. The value $n = 2$ leads to R in the millimetre regime, which is largely disfavoured by existing experimental data. But for $n = 3$ one arrives at the nanometre scale, which is in no conflict with any existing experiment.

The above reformulation of the hierarchy problem thus opens up the possibility of observing extra dimensions, either through scattering experiments at colliders or through sub-millimetre tests of Newton's law; see Rubakov (2001) and the references therein. Higher-dimensional theories generically predict a violation of the Newtonian $1/r$ potential at some scale, which could be seen in laboratory experiments. Another possibility is to look for signatures of the D-dimensional graviton. According to (9.75), this graviton can be recast as a tower of states (called the Kaluza–Klein or KK tower) with increasing mass. Such states could be detectable if m_D were in the TeV energy range. A search in this direction is being performed, for example, at the Large Hadron Collider (LHC). So far, no such signatures have been found.[8]

[6] Often, a reduced D-dimensional Planck mass M_D is introduced via $M_D^{2+n} = (8\pi G_D)^{-1}$.

[7] It is often assumed that the extra dimensions are compactified as a torus; then, $V_n = (2\pi R)^n$.

[8] A review of the experimental status can be found in Nakamura and Particle Data Group (2010).

In both the traditional Kaluza–Klein and the ADD scenario, the full metric factorizes[9] into a four-dimensional part describing our macroscopic dimensions and a (compact) part referring to the extra dimensions. Such an assumption is, however, not obligatory. If factorization does not hold, one talks about a ‘warped metric’. The extra dimensions can be compact or infinite in size. A warped metric occurs, for example, if the gravitational field produced by the brane is taken into account. We shall briefly describe one particular model with two branes put forward by Randall and Sundrum (1999). There is one extra dimension, and the bulk has an anti-de Sitter (AdS) geometry. All four-dimensional slices have the same geometry (they are flat), but the higher-dimensional space–time is curved.

The action of this model is given by the following expression, in which the index $I = \pm$ labels the two branes with tensions $\sigma_\pm$:

$$\begin{aligned} S[G, g, \phi] &= S_5[G] + \sum_I \int_{\Sigma_I} \mathrm{d}^4x \, \left(L_\mathrm{m}(\phi, \partial\phi, g) - g^{1/2}\sigma_I + \frac{1}{8\pi G_5}[K] \right), \\ S_5[G] &= \frac{1}{16\pi G_5} \int\limits_{M^5} \mathrm{d}^5x \; G^{1/2} \left({}^{(5)}R(G) - 2\Lambda_5 \right). \end{aligned} \tag{9.80}$$

Here, $S[G, g, \phi]$ is the action of the five-dimensional gravitational field with the metric $G = G_{AB}(x, y)$, $A = (\mu, 5)$, $\mu = 0, 1, 2, 3$, propagating in the bulk space–time ($x^A = (x, y)$, $x = x^\mu$, $x^5 = y$); and the matter fields ϕ are confined to the branes Σ_I, which are four-dimensional time-like surfaces embedded in the bulk. The branes carry the induced metrics $g = g_{\mu\nu}(x)$ and the matter field Lagrangians $L_\mathrm{m}(\phi, \partial\phi, g)$. The bulk part of the action contains the five-dimensional gravitational and cosmological constants, G_5 and Λ_5, while the brane parts have four-dimensional cosmological constants σ_I. The bulk cosmological constant Λ_5 is negative and is thus capable of generating an AdS geometry, while the brane cosmological constants play the role of brane tensions σ_I and, depending on the model, can be of either sign. The Einstein–Hilbert bulk action in (9.80) is accompanied by brane surface terms (cf. (1.1)) containing the jump of the extrinsic curvature trace $[K]$ across the two branes.

The fifth dimension has the topology of a circle with points labelled by the coordinate y, $-d < y \le d$, with an orbifold $\mathbb{Z}_2$-identification of points y and $-y$.[10] The branes Σ_+ and Σ_- are located at antipodal fixed points of the orbifold, $y = y_\pm$, where $y_+ = 0$ and $|y_-| = d$, respectively; see Fig. 9.2. When they are empty, that is, $L_\mathrm{m}(\phi, \partial\phi, g_{\mu\nu}) = 0$, and their tensions are opposite in sign and fine-tuned to the values of Λ_5 and G_5, that is,

$$\Lambda_5 = -\frac{6}{l^2}\,, \quad \sigma_+ = -\sigma_- = \frac{3}{4\pi G_5 l}, \tag{9.81}$$

this model admits a solution with an AdS metric in the bulk (l is its curvature radius),

[9] Factorization here means that the D-dimensional metric can be put into block-diagonal form, in which one block is the four-dimensional metric, and the various blocks do not depend on the coordinates referring to the other blocks.

[10] An orbifold is a coset space $M/\mathcal{G}$, where $\mathcal{G}$ is a group of discrete symmetries of the manifold M. Here we deal with the special case of an $S^1/\mathbb{Z}_2$ orbifold.

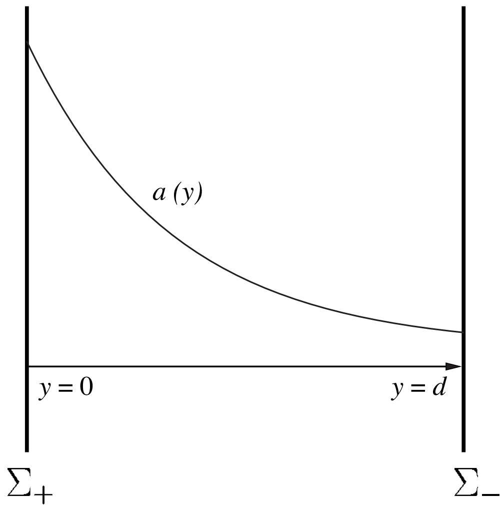

Fig. 9.2 The situation in the Randall–Sundrum model with two branes. The warp factor is $a(y) = \exp(-|y|/l)$.

$$ds^2 = dy^2 + e^{-2|y|/l}\eta_{\mu\nu}\,dx^\mu\,dx^\nu, \tag{9.82}$$

$0 = y_+ \leq |y| \leq y_- = d$, and with a flat induced metric $\eta_{\mu\nu}$ on both branes. The metric on the negative-tension brane is rescaled by the 'warp factor' $\exp(-2d/l)$.

The Randall–Sundrum scenario could provide an alternative solution to the hierarchy problem. On the first brane (the 'Planck brane' Σ_+), gravity is strong, while it becomes weak on the other brane (the 'Standard Model brane' Σ_-) due to the warp factor $\exp(-|y|/l)$. Along the fifth dimension, the Planck mass is thus modified according to $m_P\exp(-d/l)$ and can thus reach the desired TeV scale for the rather 'normal' value $d/l \approx 37$. Similarly to the case of the ADD model, one can look for experimental signatures of the Randall–Sundrum scenario at the LHC and elsewhere (Nakamura and Particle Data Group 2010).

With the fine tuning (9.81), the solution (9.82) exists for an arbitrary brane separation d—the two flat branes stay in equilibrium. Their flatness is the result of a compensation between the bulk cosmological constant and the brane tensions. We note that the distance between the branes need not be large, unlike the ADD scenario. This is because of the exponential decrease of the gravitational 'force' between the branes.

More generally, one can consider the Randall–Sundrum model with weak matter sources for metric perturbations $h_{AB}(x,y)$ on the background of this solution,

$$ds^2 = dy^2 + e^{-2|y|/l}\eta_{\mu\nu}\,dx^\mu\,dx^\nu + h_{AB}(x,y)\,dx^A\,dx^B, \tag{9.83}$$

such that this five-dimensional metric *induces* on the branes two four-dimensional metrics of the form

$$g^{\pm}_{\mu\nu}(x) = a^2_\pm\,\eta_{\mu\nu} + h^{\pm}_{\mu\nu}(x). \tag{9.84}$$

Here the scale factors $a_\pm = a(y_\pm)$ can be expressed in terms of the interbrane distance,

$$a_+ = 1, \;\; a_- = \mathrm{e}^{-2d/l} =: a, \tag{9.85}$$

and $h^\pm_{\mu\nu}(x)$ are the perturbations by which the brane metrics $g^\pm_{\mu\nu}(x)$ differ from the (conformally) flat metric in the Randall–Sundrum solution (9.82).

Instead of using the Kaluza–Klein formalism with its infinite tower of modes, one can employ an alternative formalism that captures more the spirit of the holographic principle and the AdS/CFT correspondence (Section 9.2.8). This results in the calculation of a non-local effective action for the branes; see Barvinsky *et al.* (2003*a*) for details. This action is a functional of the induced metrics on the two branes. One can also derive a reduced action which depends only on one brane (the 'visible', that is, 'our' brane, which is taken here to be Σ_+). These effective actions contain *all* the physical information that is available on the brane(s). This has interesting applications in inflationary cosmology and gravitational-wave interferometry. The latter arises because this model has *light* massive graviton modes in addition to the massless graviton. The mixing of these modes can lead to gravitational-wave oscillations analogous to neutrino oscillations. The parameters of these oscillations depend crucially on the size of the extra dimension. Such an effect can, in principle, be observed with current gravitational-wave interferometers. The mechanism of this mixing and its phenomenology are discussed in Barvinsky *et al.* (2003*b*), to which we refer the reader for more details.

9.2.7 On the formation of small black holes in accelerators

In the last subsection, we discussed the hierarchy problem of the Standard Model and its possible solution by the introduction of extra dimensions. If this were the solution, the fundamental Planck scale would be comparable to the electroweak scale, that is, $m_D \sim 1$ TeV. But then there would be the exciting possibility of actually producing small black holes in accelerators and thus probing effects of quantum gravity by their evaporation (Dimopoulos and Landsberg 2001, Giddings and Thomas 2002). The TeV energy range can be reached by the Large Hadron Collider at CERN in Geneva. This is a storage ring which works with proton–proton collisions at a centre-of-mass energy of eventually $\sqrt{s} = 14$ TeV. For this reason, the LHC is well suited to probing the physics at the TeV scale.

One expects that black holes will be produced in these collisions from energies of the order of m_D on. A black hole, as such, is a classical object characterized by an event horizon. But for energies around m_D, one would not expect that such a classical object would form. Instead, such an object should exhibit genuine quantum properties and is thus called a 'quantum black hole'. Its properties can, however, be inferred only from a full quantum theory of gravity. Quantum black holes are not expected to decay thermally. The experimental study of these objects could thus provide important clues in the search for such a theory. In contrast, black holes with masses $M \gg m_D$ should be characterizable as usual by an event horizon, and are thus referred to as 'semiclassical black holes'. One expects that they will decay thermally via Hawking radiation. Many investigations assume that the scenario developed for semiclassical black holes will continue to apply at least approximately as one approaches m_D from above.

Let us discuss black-hole formation in the context of the ADD model (see Section 9.2.6).[11] We restrict ourselves here to Schwarzschild black holes, although one would assume that a black hole with angular momentum will be formed in a generic collision. If the Schwarzschild radius is much smaller than the compactification radius R, one can use the expression for the Schwarzschild radius $R_{\rm S}$ of a black hole embedded in $3+n$ flat spatial dimensions as derived by Myers and Perry (1986):

$$R_{\rm S}(M) = \frac{1}{\sqrt{\pi} m_D} \left(\frac{M}{m_D} \frac{8\Gamma\left((n+3)/2\right)}{n+2} \right)^{1/(n+1)} .$$

Assuming $M \sim m_D$, $R_{\rm S}(M)$ will be of the order of the Compton wavelength corresponding to m_D. For example, for $n = 3$ and $m_D \sim 1$ TeV, $R_{\rm S}$ will be of order 10^{-19} m, which is definitely much smaller than the compactification radius for $n = 3$ in the ADD model, which is in the nanometre regime; the assumption $R_{\rm S} \ll R$ is thus well satisfied.

Concrete calculations lead to the result that the cross-section for black-hole formation is of the order of the geometrical cross-section (Dimopoulos and Landsberg 2001):

$$\sigma(M) \approx \pi R_{\rm S}^2 .$$

At LHC energies, this is of the order of 100 pb if $m_D \sim 1$ TeV.[12] Because there is no suppression of black-hole formation by a small dimensionless coupling constant such as the fine-structure constant, black holes would be produced at enormous rates once the necessary energy was available.[13] If such black holes exist, the probability of seeing them at the LHC would be overwhelming.

Once produced, these semiclassical black holes should decay via Hawking radiation (Section 7.1.2). This decay should happen with equal probabilities ('democratically') to all degrees of freedom of the Standard Model, since the gravitational coupling is not sensitive to flavour. Most of the emitted particles would be quarks and gluons (about 75%), because they have a large number of colour degrees of freedom. Only a tiny fraction of the particles would be gravitons and neutrinos; this is advantageous because these particles are undetectable, and so the whole process is characterized by little missing energy.

The Hawking temperature for such a higher-dimensional black hole is (Myers and Perry 1986)

$$T_{\rm BH} = \frac{\hbar(n+1)}{4\pi R_{\rm S}}, \tag{9.86}$$

which for $n = 0$ coincides, of course, with the four-dimensional expression (1.35). One can estimate the lifetime of such a black hole using similar arguments to those in

[11] Similar investigations can be done for the Randall–Sundrum model.

[12] 1 barn $\equiv$ 1 b $= 10^{-24}$ cm^2.

[13] Dimopoulos and Landsberg (2001) give a production rate of one per second!

Section 7.1.2 (see (7.17)). The evaporation occurs primarily in three spatial dimensions and thus

$$\frac{\mathrm{d}M}{\mathrm{d}t} \propto R_{\mathrm{S}}^2 T_{\mathrm{BH}}^4 \propto R_{\mathrm{S}}^{-2},$$

from which one gets the following estimate for the lifetime (inserting all units):

$$\tau_{\mathrm{BH}} \sim \frac{\hbar}{m_D c^2}\left(\frac{M}{m_D}\right)^{(n+3)/(n+1)},$$

which is about 10^{-27} s for $m_D \sim 1$ TeV.

Under the assumption that the black holes produced can be treated semiclassically, the CMS Collaboration has already published experimental limits (CMS Collaboration 2011). In the framework of the ADD model, these limits exclude the production of black holes with a minimum mass in the range 3.5–4.5 TeV for values of m_D up to 3 TeV at 95% C.L.

It should be possible to check the Hawking temperature (9.86) experimentally by determining M from the energy of the decay products and T_{BH} from the energy spectrum of the emitted particles (Dimopoulos and Landsberg 2001, Kanti 2004). One could then also infer the number of extra dimensions from these measurements. Another interesting feature would be the possible generation of new particles with mass ~ 100 GeV (e.g. a light Higgs particle).

If such small black holes were discovered at the LHC, it should also be possible to observe them in cosmic rays, for example at the Pierre Auger Observatory in Argentina. They would then also play a role as primordial black holes (Section 7.7).

As mentioned above, black holes with a mass $M \sim m_D$ should behave in a fully quantum way (see e.g. Calmet 2010). It is imaginable that only a small number of semiclassical black holes will be formed and that these quantum black holes will therefore be the dominant objects. They will decay non-thermally and should exhibit themselves dominantly by the emission of two jets (Calmet 2010). The ATLAS Collaboration at the LHC was able to give limits on quantum-black-hole production in the context of concrete models (ATLAS Collaboration 2011). It is thus currently open whether such small black holes and their underlying extra dimensions really exist.

9.2.8 Holographic principle and AdS/CFT correspondence

The Bekenstein–Hawking entropy is calculated using the surface area of the horizon, not the volume inside. The idea that for a gravitating system the information is located on the *boundary* of some spatial region is called the 'holographic principle' (see Bousso (2002) for a review). More generally, this principle states that the number of degrees of freedom in the volume of a spatial region is equal to that of a system residing on the boundary of that region. Evidently, this gives rise to non-local features.

The holographic principle seems to be realized in string theory in the 'AdS/CFT correspondence'; see Aharony *et al.* (2000) and Maldacena (2011) for reviews. This is the second approach, in addition to matrix theory, by which we can learn something about M-theory. The AdS/CFT correspondence states that non-perturbative string theory in a background space–time that is asymptotically anti-de Sitter (AdS) is dual to a conformal field theory (CFT) defined in a flat space–time of one dimension fewer.

Type IIB string theory on an asymptotically $\mathrm{AdS}_5 \times S^5$ space–time (called the 'bulk') may serve as a concrete example. This theory is dual to a CFT that is a (3+1)-dimensional SUSY Yang–Mills theory with gauge group U(n). Since the conformal boundary of AdS_5 is $\mathbb{R} \times S^3$, whose dimension agrees with that of the CFT, it is claimed that the CFT is defined on the boundary of AdS space. This cannot be meant literally, since a boundary cannot in general be separated from the enclosed volume, because of quantum entanglement between the two. It should also be mentioned that in a more recent version of the AdS/CFT correspondence, genuine string calculations can be compared with calculations in gauge theories; cf. Berenstein *et al.* (2002). A possible correspondence between a two-dimensional CFT and an extremal Kerr black hole was conjectured by Guica *et al.* (2009).

The AdS/CFT correspondence (also called the 'AdS/CFT conjecture') associates fields in string theory with operators in the CFT and compares expectation values and symmetries in the two theories.[14] An equivalence at the level of the quantum states in the two theories has not been shown; see Giddings (2011). As long as no independent non-perturbative definition of string theory exists, the AdS/CFT conjecture cannot be proven. This correspondence is often interpreted as a non-perturbative and mostly background-independent definition of string theory, since the CFT is defined non-perturbatively and the background metric enters only through boundary conditions at infinity; cf. Horowitz (2005) and Blau and Theisen (2009). Full background independence in the sense of canonical quantum gravity has, however, not yet been implemented. A major open question is, of course, 'what are the fundamental symmetries underlying string theory?'

9.3 String field theory

In Section 3.2, we treated the bosonic string in the functional Schrödinger picture. The string was described there by a wave functional $\Psi[X^\mu(\sigma)]$, where the $X^\mu(\sigma)$ are the embedding variables with respect to the D-dimensional space–time in which the string propagates. This functional is also called a 'string field'. For the closed string, it can be expanded in momentum space as (see e.g. Taylor 2009)

$$\Psi = \int \mathrm{d}^D p \ \left[\phi(p)|p\rangle + g_{\mu\nu}(\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1} + \alpha^\nu_{-1}\tilde{\alpha}^\mu_{-1}|p\rangle) + \ldots\right] . \qquad (9.87)$$

The sum includes the dilaton ϕ, the graviton as described by the metric $g_{\mu\nu}$, and other massless states and contributions from the infinitely many excited (massive) string states. The string field thus contains an infinite number of space–time fields.

In ordinary quantum field theory, the fields $\psi(x^\mu)$ are turned into operators. In string field theory, the functionals $\Psi[X^\mu(\sigma)]$ are turned into operators. One defines for this purpose an action functional $S[\Psi]$, which is then used as an ingredient for a path integral where a sum over all fields Ψ is performed; this corresponds to a path integral over the infinite set of fields contained in (9.87). Such a path integral is, of

[14] As Horowitz and Polchinski in Oriti (2009, p. 230) put it, 'AdS/CFT and other dualities are statements about mathematical physics, which can be used to derive relations between the spectrum, amplitudes, and other physical properties of the two sides of the duality.'

course, extremely difficult to deal with, since it contains infinitely many fields with complicated (partly non-local) interactions. As argued, for example, by Taylor (2009), string field theory is currently the only truly background-independent approach to string theory. This argument is supported by the observation that, at least for open strings, different backgrounds can be accommodated into the same open-string field theory (Sen 1999).

For quantum gravity, we need closed strings, but for reasons of simplicity we shall briefly mention Witten's open-bosonic-string field theory (Witten 1986) here. In analogy to (9.87), the string field can be expanded in the form

$$\Psi = \int \mathrm{d}^{26}p \ \left[\varphi(p)|p\rangle \ldots\right] \ ;$$

because we are dealing with open strings here, the metric does not occur.

Witten has proposed an action that has the form of a Chern–Simons action on a three-dimensional manifold, although this proposal does not relate it to three spatial dimensions. The action can be written as the sum of a free term and an interaction term,

$$S[\Psi] = \frac{1}{2}\int \Psi \star Q\Psi + \frac{2g_{\mathrm{o}}}{3}\int \Psi \star \Psi \star \Psi.$$

Here, $\star$ denotes a product operation for the string fields, and g_{o} is the coupling constant for the open string; Q is a BRST operator.

It has been argued that one could use string field theory to describe the string landscape, which is the set of the (10^{500} or more) minima of the (largely unknown) string effective potential (cf. Taylor 2009). This is supported by the results of Sen (1999), who found for the effective tachyon potential in open-string field theory a new stable vacuum below the perturbative unstable vacuum that was known before. This shows how distinct vacua and thus perhaps the full string landscape might arise from an effective potential in string field theory.

Closed strings are implicitly contained in the theory of open strings. They should thus also appear in the open-string field theory sketched above. However, because interactions at all orders are involved, the direct construction of a closed-string field theory seems too difficult to achieve at present.

Is there any relation of the string field to the Wheeler–DeWitt wave functional discussed in Chapter 5? Since the metric occurs in the expansion (9.87), one might expect that at the level of canonical quantization there would be a close resemblance of the string field Ψ to the Wheeler–DeWitt functional, which depends on the three-metric and on non-gravitational fields. Turning the string field into an operator would correspond to a 'third quantization' of the Wheeler–DeWitt equation (McGuigan 1988); cf. also Section 5.2.2. The meaning of such a third quantization is, however, presently open.

Further reading: Green *et al.* (1987), Polchinski (1998*a*,*b*), Zwiebach (2009).

10 Phenomenology, decoherence, and the arrow of time

10.1 Quantum-gravity phenomenology

A central feature of any physical theory is its testability by experiment. The same applies to quantum theory. As long as no tests exist, we shall stay with the situation of having many approaches at our disposal, among which no discrimination can be made. It is therefore of great importance to devise experiments or observations that could at least give some hints about the direction in which we could look for the correct theory. This is the field of quantum-gravity phenomenology. Although it has some similarities to particle-physics phenomenology, the situation is not quite the same. Much work in quantum-gravity phenomenology is devoted not so much to calculating concrete predictions from specific approaches (which is often too difficult) as to looking for violations of established laws and principles that *could* arise from a quantum-gravity effect.

Nevertheless, there exist some cases where concrete predictions from specific approaches have been made. We have already encountered several examples in the course of this book. Atomic transition rates through graviton emission can be calculated, although these rates turn out to be too small to be observable (Section 2.1.3). The same applies for the quantum-gravitational corrections to the Newtonian potential discussed in Section 2.2.3. Similar effects can also be calculated from the Wheeler–DeWitt equation and from loop quantum cosmology. One can derive, for example, quantum-gravitational modifications of the CMB anisotropy spectrum (Sections 5.4 and 8.4). If primordial black holes (Section 7.7) or the production of small black holes at accelerators (Section 9.2.7) were discovered, they could also provide signatures of quantum gravity. From string theory (Chapter 9), it should be possible to predict low-energy coupling constants and parameters, but nothing in this direction has been achieved so far; one might even find that predictions can only be made on the basis of the anthropic principle (Section 10.3).

Effects of quantum gravity may also be seen in the CMB spectrum in a different sense. Inflationary models of the early universe predict the production of gravitons. It must be emphasized that this production by cosmological evolution is an effect of linear quantum gravity. This important effect was discussed by Starobinsky (1979). Future experiments may be able to see the contribution of those gravitons to the CMB spectrum. It is even possible that they may be observed directly in future missions searching for gravitational waves in space. Moreover, a possible discreteness in the inflationary perturbations could manifest itself in the spectrum (Hogan 2002).

Here we shall be concerned mainly with the search for violations of established laws. One possibility is the violation of Lorentz invariance. Such a violation may originate from a discrete structure of space–time, for which evidence is found in string theory, loop quantum gravity, and other approaches. Although Lorentz invariance is not a symmetry of GR, it is valid there in a local inertial system; it is, in fact, part of the Einstein equivalence principle. When we talk about violation, we thus refer to local violation in a freely falling system. A violation of Lorentz invariance would signal the existence of a preferred frame of reference. But where would such a preferred system (an 'aether') come from? Formally, one can describe it by, for example, coupling a time-like vector field to geometric quantities.

In a pure phenomenological approach, one makes an ad hoc ansatz for a modified dispersion relation without trying to answer the question of where the modification comes from. That is, one makes an ansatz of the form (restoring c)

$$E^2 = p^2c^2 + m^2c^4 + f(E, \mathbf{p}, m_{\rm P}), \tag{10.1}$$

where, instead of the Planck mass, one could also insert another (so far unknown) mass scale κ. The unknown function f is expanded in powers of $E/m_{\rm P}c^2$, and attention is then focused on the various terms of this expansion. It was noted in Amelino-Camelia *et al.* (1998) that γ-ray bursts could be sensitive to a modified dispersion relation (γ-ray bursts are short, energetic pulses that mostly originate at cosmological distances).

Violation of Lorentz invariance is often discussed in connection with non-commutative space–times. There, the coordinates are non-commutative,

$$[x_\mu, x_\nu] = \mathrm{i}\,\Theta_{\mu\nu} + \mathrm{i}\,\rho^\alpha_{\mu\nu} x_\alpha, \tag{10.2}$$

which in the general case leads to a violation of Lorentz invariance. There are exceptions such as κ-Minkowski non-commutative space–time, given by

$$[x_k, t] = \frac{\mathrm{i}\hbar}{\kappa c^2} x_k \ , \qquad [x_k, x_l] = 0,$$

where κ is a new distinguished mass scale.

Violation of Lorentz invariance may also be accompanied by a violation of CPT symmetry. After all, one of the main assumptions of the CPT theorem in quantum field theory is Lorentz invariance. So far, no violations of Lorentz or CPT invariance have been seen; Kostelecký and Russell (2011) have presented detailed tables showing the empirical limits for such a violation.

An alternative to the violation of Lorentz invariance is a modification of Lorentz invariance that is in accordance with the principle of relativity, that is, which does not introduce a preferred frame. The result is 'deformed' (or 'doubly') special relativity (DSR). Here, instead of just c as in special relativity, one has *two* invariants: c and $m_{\rm P}$ (or another mass scale κ). This leads again to modified dispersion relations as in (10.1). DSR may also be connected to the κ-Minkowski non-commutative space–time mentioned above.

Unfortunately, DSR faces some conceptual problems, such as the difficulty of obtaining a correct description of multi-particle states. Up to now, there has been no empirical evidence for DSR. For example, the results of the search for energy-dependent arrival times of high-energy photons from γ-ray bursts have been negative (Abdo *et al.* 2009).

Besides violation of Lorentz invariance, a new fundamental theory may also lead to violation of the universality of free fall and the gravitational redshift. One may also look for signatures of 'fluctuations' at the Planck scale. They might be seen with the help of atomic interferometry (Percival and Strunz 1997), although there exist strong observational constraints (Peters *et al.* 2001). A detailed treatment of quantum-gravity phenomenology can be found in the book edited by Amelino-Camelia and Kowalski-Glikman (2005); shorter introductions and reviews include Hossenfelder (2010) and Amelino-Camelia *et al.* (2005).

10.2 Decoherence and the quantum universe

The central concept in quantum theory is the superposition principle. It consists of a kinematical and a dynamical part; cf. Joos *et al.* (2003). The kinematical part declares that for any two physical states Ψ_1 and Ψ_2, the sum $c_1\Psi_1 + c_2\Psi_2$ (with $c_1, c_2 \in \mathbb{C}$) is again a physical state. This expresses the linear structure of Hilbert space and gives rise to the important notion of quantum *entanglement* between systems. The dynamical part refers to the linearity of the Schrödinger equation. If $\Psi_1(t)$ and $\Psi_2(t)$ are solutions, then the sum $c_1\Psi_1(t) + c_2\Psi_2(t)$ is again a solution. The superposition principle has so far passed all experimental tests; cf. Schlosshauer (2006).

The superposition principle remains untouched in most approaches to quantum gravity. This holds in particular for quantum GR and string theory, which are both discussed in this book. There have been suggestions that a gravity-induced breakdown of the superposition principle occurs; see for example, Penrose (1996) and the discussion in Chapter 8 of Joos *et al.* (2003). However, no such mechanism has been developed to a technical level comparable with the approaches discussed here.

If the superposition principle is universally valid, quantum gravity allows the superposition of macroscopically different metrics. This has, in particular, drastic consequences for quantum cosmology (Chapter 8), where it would be difficult to understand why we observe a classical universe at all. In some interpretations of quantum mechanics, notably the Copenhagen interpretation(s), an external observer who 'reduces' the wave function from superpositions to the observed component is invoked. In quantum cosmology, on the other hand, no such external measuring agency is available, since the Universe contains by definition everything. A reduction (or collapse) of the wave function by external observers is then impossible. How, then, does the classical appearance of our Universe emerge? In recent decades, an understanding has been reached of how classical properties can emerge within quantum mechanics. It is an amazing fact that the key role in this process is played by the superposition principle itself, through the process of *decoherence*. Following Kiefer (2003*b*), we shall give a brief introduction to decoherence in quantum mechanics and then extrapolate decoherence into the realm of quantum cosmology (Section 10.1.2). An exhaustive treatment can be found in Joos *et al.* (2003); see also Zurek (2003) and Schlosshauer (2007).

10.2.1 Decoherence in quantum mechanics

If quantum theory is universally valid, every system should be described in quantum terms, and it would be inconsistent to draw an a priori border line between a quantum system and a classical apparatus. The measurement process was analysed within quantum mechanics for the first time by von Neumann (1932). He considered the coupling of a system (S) to an apparatus (A) (see Fig. 10.1).

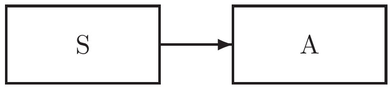

Fig. 10.1 Original form of the von Neumann measurement model.

If the states of the measured system that are to be distinguished by the apparatus are denoted by $|n\rangle$ (e.g. spin up and spin down), an appropriate interaction Hamiltonian has the form

$$H_{\text{int}} = \sum_n |n\rangle\langle n| \otimes \hat{A}_n. \tag{10.3}$$

The operators $\hat{A}_n$ act on the states of the apparatus and are rather arbitrary, but must, of course, depend on the 'quantum number' n. Equation (10.3) describes an 'ideal' interaction, during which the apparatus becomes correlated with the system state without changing the latter. There is thus no disturbance of the system by the apparatus—on the contrary, the apparatus is disturbed by the system (in order to yield a measurement result).

If the measured system is initially in the state $|n\rangle$ and the device in some initial state $|\Phi_0\rangle$, the evolution according to the Schrödinger equation with the Hamiltonian (10.3) reads

$$\begin{aligned} |n\rangle|\Phi_0\rangle \stackrel{t}{\longrightarrow} \exp\left(-\mathrm{i}H_{\text{int}}t\right)|n\rangle|\Phi_0\rangle &= |n\rangle \exp\left(-\mathrm{i}\hat{A}_n t\right)|\Phi_0\rangle \\ &=: |n\rangle|\Phi_n(t)\rangle. \end{aligned} \tag{10.4}$$

The resulting apparatus states $|\Phi_n(t)\rangle$ are often called 'pointer states'. A process analogous to (10.4) can also be formulated in classical physics. The essential new quantum features come into play when one considers a *superposition* of different eigenstates (of the measured 'observable') as the initial state. The linearity of time evolution immediately leads to

$$\left(\sum_n c_n|n\rangle\right)|\Phi_0\rangle \stackrel{t}{\longrightarrow} \sum_n c_n|n\rangle|\Phi_n(t)\rangle. \tag{10.5}$$

But this state is a superposition of macroscopic measurement results (of which Schrödinger's cat is just one drastic example)! To avoid such a bizarre state, and to avoid the apparent conflict with experience, von Neumann introduced a dynamical collapse of the wave function as a new law that violates the universality of the Schrödinger equation. The collapse should then select one component with probability $|c_n|^2$.

Can von Neumann's conclusion and the introduction of a collapse be avoided? The crucial observation that enforces an extension of von Neumann's measurement theory is the fact that macroscopic objects (such as measurement devices) are so strongly coupled to their natural environment that a unitary treatment as in (10.4) is by no means sufficient and has to be modified to include the environment (Zeh 1970); see Fig. 10.2.

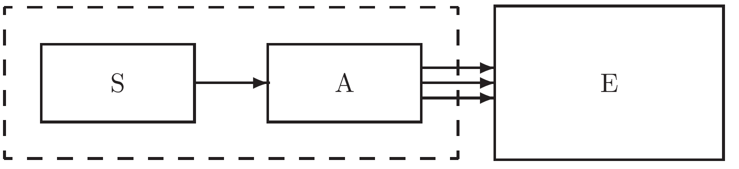

Fig. 10.2 Realistic extension of the von Neumann measurement model including the environment. Classical properties emerge through the unavoidable, irreversible interaction of the apparatus with the environment.

Fortunately, this can easily be done to a good approximation, since the interaction with the environment has in many situations the same form as that given by the Hamiltonian (10.3): the measurement device is itself 'measured' (passively recognized) by the environment, according to

$$\left(\sum_n c_n|n\rangle|\Phi_n\rangle\right)|E_0\rangle \quad \stackrel{t}{\longrightarrow} \quad \sum_n c_n|n\rangle|\Phi_n\rangle|E_n\rangle. \tag{10.6}$$

This is again a macroscopic superposition, now including the myriads of degrees of freedom pertaining to the environment (gas molecules, photons, etc.). However, most of these environmental degrees of freedom are inaccessible. Therefore, they have to be integrated out from the full state (10.6). This leads to a reduced density matrix for the system plus the apparatus, which contains all the information that is available there. It reads

$$\rho_{\mathrm{SA}} \approx \sum_n |c_n|^2|n\rangle\langle n| \otimes |\Phi_n\rangle\langle\Phi_n| \qquad \text{if} \qquad \langle E_n|E_m\rangle \approx \delta_{nm}, \tag{10.7}$$

since, under realistic conditions, different environmental states are orthogonal to each other. Equation (10.7) is identical to the density matrix of an ensemble of measurement results $|n\rangle|\Phi_n\rangle$. The system and apparatus thus seem to be in one of the states $|n\rangle$ and $|\Phi_n\rangle$, given by the probability $|c_n|^2$.

Both system and apparatus thus assume classical properties through the unavoidable, irreversible interaction with the environment. This dynamical process, which is fully described by quantum theory, is called decoherence. It is based on the quantum entanglement between the apparatus and the environment. In ordinary macroscopic situations, decoherence occurs on an extremely short timescale, giving the impression of an instantaneous collapse or a 'quantum jump'. Recent experiments have been able to demonstrate the continuous emergence of classical properties in mesoscopic systems (Hornberger *et al.* 2003, Schlosshauer 2007). Therefore, one would never be able to observe a weird superposition such as Schrödinger's cat, because the information about this superposition would be delocalized almost instantaneously into unobservable correlations with the environment, resulting in an *apparent* collapse for the cat state.

The interaction with the environment distinguishes the local basis with respect to which classical properties (unobservability of interferences) hold. This 'pointer basis' must obey the condition of robustness; that is, it must keep its classical appearance over the relevant timescales; cf. Zurek (2003). Classical properties are thus not intrinsic to any object, but only defined by their interaction with other degrees of freedom. In simple (Markovian, i.e. local in time) situations, the pointer states are given by localized Gaussian states (Diósi and Kiefer 2000). They are, in particular, relevant to the localization of macroscopic objects.

The ubiquitous occurrence of decoherence renders the interpretational problem of quantum theory at present largely a 'matter of taste' (Zeh 1994). Provided one adopts a realistic interpretation without additional variables,[1] the alternatives would be to have either an Everett interpretation or the assumption of a collapse for the total system (including the environment). The latter would have to entail an explicit modification of quantum theory, since one would have to introduce non-linear or stochastic terms into the Schrödinger equation in order to achieve this goal. The Everett interpretation assumes that all components of the full quantum state exist and are real. Decoherence produces robust macroscopic branches, one of which corresponds to the observed world. Interference with the other branches is suppressed, so decoherence readily explains the observation of an apparent collapse of the wave function, independent of whether there is a real collapse for the total system or not. The question is thus whether one applies Ockham's razor[2] to the equations or the intuition (Zeh 1994): either one has to complicate the formalism in order to have just one macroscopic branch or one retains the linear structure of quantum theory and has to accept the existence of 'many worlds'.

10.2.2 Decoherence in quantum cosmology

In this subsection, we investigate the following question: how can one understand the classical appearance of global space–time variables such as the radius (scale factor) of the universe? If decoherence is the fundamental process, we have to identify a 'system' and an 'environment'. More precisely, we have to differentiate between relevant ('accessible') and irrelevant ('inaccessible') variables. All degrees of freedom exist, of course, within the universe. It has been suggested by Zeh (1986) that the irrelevant degrees of freedom are the variables describing density fluctuations and gravitational waves. Their interaction with the scale factor and other homogeneous degrees of freedom (such as an inflaton field) can render the latter classical. In a sense, then, a classical space–time arises from a 'self-measurement' of the universe.

The following discussion will roughly follow Joos *et al.* (2003). Calculations of decoherence in quantum cosmology can be performed in the framework of quantum geometrodynamics (Kiefer 1987), using the formalism of the Wheeler–DeWitt equation presented in Chapters 5 and 8. A prerequisite is the validity of the semiclassical

[1] The Bohm theory would be an example of a realistic approach with additional variables.

[2] 'Pluralitas non est ponenda sine necessitate'.

approximation (Section 5.4) for the global variables. This brings an approximate time parameter t into play. The irrelevant degrees of freedom (density fluctuations, gravitational waves) are described by the inhomogeneous variables of Section 8.2. In Kiefer (1987), the relevant system was taken to be the scale factor ('radius') a of the universe together with a homogeneous scalar field ϕ (the 'inflaton'); cf. the model discussed in Section 8.1.2. The inhomogeneous modes of Section 8.2 can then be shown to *decohere* the global variables a and φ.

An open problem in Kiefer (1987) was the issue of regularization; the number of fluctuations is infinite and would cause divergences, so an ad hoc cut-off was suggested, only modes with a wavelength greater than the Planck length being retained. The problem was again addressed in Kiefer (1992) and in more detail in Barvinsky *et al.* (1999*a*), where a physically motivated regularization scheme was introduced. In the following, we shall briefly review this approach.

As a (semi)classical solution for a and ϕ, one may use

$$\phi(t) \approx \phi, \tag{10.8}$$

$$a(t) \approx \frac{1}{H(\phi)} \cosh H(\phi)t, \tag{10.9}$$

where $H^2(\phi) = 4\pi V(\phi)/3m_{\mathrm{P}}^2$ is the Hubble parameter generated by the inflaton potential $V(\phi)$; cf. Section 8.3.2. It is approximately constant during the inflationary phase, in which ϕ slowly 'rolls down' the potential. We take into account fluctuations of a field $f(t,\mathbf{x})$, which can be a field of any spin (not necessarily a scalar field Φ). Space is assumed to be a closed three-sphere, so $f(t,\mathbf{x})$ can be expanded in a discrete series of orthonormal spherical harmonics $Q^n(\mathbf{x})$ (cf. Section 8.2),

$$f(t,\mathbf{x}) = \sum_{\{n\}} f_n(t) Q^n(\mathbf{x}). \tag{10.10}$$

One can thus represent the fluctuations by the degrees of freedom f_n (in Section 8.2, the f_n were the modes of a scalar field Φ).

The aim is now to solve the Wheeler–DeWitt equation in the semiclassical approximation. This leads to the solution

$$\Psi(t|\phi, f) = \frac{1}{\sqrt{v_\phi^*(t)}}\, \mathrm{e}^{-I(\phi)/2 + \mathrm{i}S_0(t,\phi)} \prod_n \psi_n(t,\phi|f_n). \tag{10.11}$$

The time t that appears here is the semiclassical ('WKB') time and is defined by the background degrees of freedom a and ϕ through the 'eikonal' S_0, which is a solution of the Hamilton–Jacobi equation; cf. (8.50). Since ϕ is determined by a within a semiclassical branch of the wave function, only one variable (a *or* ϕ) occurs in the argument of Ψ. The wave functions ψ_n for the fluctuations f_n each obey an approximate Schrödinger equation (8.49) with respect to t, and their Hamiltonians H_n have the form of a ('time-dependent') harmonic-oscillator Hamiltonian. The first exponent in (10.11) contains the Euclidean action $I(\phi)$ from the classically forbidden region (the 'de Sitter instanton') and is independent of t. Its form depends on the boundary conditions imposed, and we shall choose here the no-boundary condition of Section 8.3.2, which amounts to $I(\phi) \approx -3m_{\rm P}^4/8V(\phi)$. The detailed form is, however, not necessary for the discussion below. The function $v_\phi(t)$ in (10.11) is the so-called basis function for ϕ and is a solution of the classical equation of motion.

For the ψ_n, we shall take the de Sitter-invariant vacuum state (the Euclidean vacuum discussed in Section 8.3.2). This reads

$$\psi_n(t,\phi|f_n) = \frac{1}{\sqrt{v_n^*(t)}} \exp\left(-\frac{1}{2}\Omega_n(t) f_n^2\right), \tag{10.12}$$

$$\Omega_n(t) = -\mathrm{i}a^3(t)\frac{\dot{v}_n^*(t)}{v_n^*(t)}. \tag{10.13}$$

The functions v_n are the basis functions of the de Sitter-invariant vacuum state; they satisfy the classical equation of motion

$$F_n\left(\frac{\mathrm{d}}{\mathrm{d}t}\right) v_n := \left(\frac{\mathrm{d}}{\mathrm{d}t}a^3\frac{\mathrm{d}}{\mathrm{d}t} + a^3m^2 + a(n^2-1)\right) v_n = 0 \tag{10.14}$$

with the boundary condition that they should correspond to a standard Minkowski positive-frequency function for constant a. In the simple special case of a spatially flat section of de Sitter space, one would have

$$a v_n = \frac{\mathrm{e}^{-in\eta}}{\sqrt{2n}}\left(1 - \frac{\mathrm{i}}{n\eta}\right), \tag{10.15}$$

where η is the conformal time defined by $a\,\mathrm{d}\eta = \mathrm{d}t$. We note that it is the corresponding *negative*-frequency function that enters the exponent of the Gaussian, see (10.13).

An important property of these vacuum states is that their norm is conserved *along any semiclassical solution* (10.8) and (10.9):

$$\left\langle \psi_n | \psi_n \right\rangle := \int \mathrm{d}f_n\, |\psi_n(f_n)|^2 = \sqrt{2\pi}[\Delta_n(\phi)]^{-1/2}, \tag{10.16}$$

$$\Delta_n(\phi) := \mathrm{i}a^3(v_n^*\dot{v}_n - \dot{v}_n^* v_n) = \text{constant}. \tag{10.17}$$

Note that $\Delta_n(\phi)$ is just the (constant) Wronskian corresponding to (10.14).[3] We must emphasize that Δ_n is a non-trivial function of the background variable ϕ, since it is defined on the full configuration space and not only along semiclassical trajectories. In a sense, it gives the weights in the 'Everett branches'. It is therefore *not* possible to normalize the ψ_n artificially to one, since this would be inconsistent from the point of view of the full Wheeler–DeWitt equation (Barvinsky *et al.* 1999*a*).

The solution (10.11) forms the basis for our discussion of decoherence. Since the $\{f_n\}$ are interpreted as the environmental degrees of freedom, they have to be integrated out to get the reduced density matrix (cf. (10.7)) for ϕ or a (a and ϕ can be

[3]The corresponding Wronskian for the homogeneous mode ϕ is $\Delta_\phi := ia^3(v_\phi^*\dot{v}_\phi - \dot{v}_\phi^* v_\phi)$.

used interchangeably, since they are connected by t). The reduced density matrix thus becomes

$$\rho(t|\phi,\phi') = \int \mathrm{d}f\ \Psi(t|\phi,f)\Psi^*(t|\phi',f), \tag{10.18}$$

where Ψ is given by (10.11), and it is understood that $\mathrm{d}f = \prod_n \mathrm{d}f_n$. After the integration, one finds

$$\begin{aligned}\rho(t|\phi,\phi') = C\frac{1}{\sqrt{v_\phi^*(t)v_\phi'(t)}}\exp\left[-\frac{1}{2}I - \frac{1}{2}I' + \mathrm{i}(S_0 - S_0')\right]\\ \times\prod_n\left[v_n^*v_n'(\Omega_n + \Omega_n'^*)\right]^{-1/2},\end{aligned} \tag{10.19}$$

where C is a numerical constant. The diagonal elements $\rho(t|\phi,\phi)$ describe the probabilities for certain values of the inflaton field to occur.

It is convenient to rewrite the expression for the density matrix (10.19) in the form

$$\begin{aligned}\rho(t|\phi,\phi') = C\frac{\Delta_\phi^{1/4}\Delta_\phi'^{1/4}}{\sqrt{v_\phi^*(t)v_\phi'(t)}}\exp\left(-\frac{1}{2}\mathbf{\Gamma} - \frac{1}{2}\mathbf{\Gamma}' + \mathrm{i}(S_0 - S_0')\right)\\ \times\boldsymbol{D}(t|\phi,\phi'),\end{aligned} \tag{10.20}$$

where

$$\mathbf{\Gamma} = I(\phi) + \mathbf{\Gamma}_{1-\mathrm{loop}}(\phi) \tag{10.21}$$

is the full Euclidean effective action, including the classical part and the one-loop part (cf. Section 2.2.4). The latter comes from the next-order WKB approximation and is important for the normalizability of the wave function with respect to ϕ. The last factor in (10.20) is the *decoherence factor*

$$\boldsymbol{D}(t|\phi,\phi') = \prod_n\left(\frac{4\,\mathrm{Re}\,\Omega_n\,\mathrm{Re}\,\Omega_n'^*}{(\Omega_n + \Omega_n'^*)^2}\right)^{1/4}\left(\frac{v_n}{v_n^*}\frac{v_n'^*}{v_n'}\right)^{1/4}. \tag{10.22}$$

It is equal to one for coinciding arguments. While the decoherence factor is time-dependent, the one-loop contribution to (10.20) does not depend on time and can play a role only at the onset of inflation. In a particular model with non-minimal coupling (Barvinsky *et al.* 1997), the magnitude of the non-diagonal elements at the onset of inflation is approximately equal to that of the diagonal elements. The universe is thus essentially quantum at this stage, that is, in a non-classical state.

The amplitude of the decoherence factor can be rewritten in the form

$$|\boldsymbol{D}(t|\phi,\phi')| = \exp\frac{1}{4}\sum_n \ln\frac{4\,\mathrm{Re}\,\Omega_n\,\mathrm{Re}\,\Omega_n'^*}{|\Omega_n + \Omega_n'^*|^2}. \tag{10.23}$$

The convergence of this series is far from guaranteed. Moreover, the divergences might not be renormalizable by local counterterms in the bare quantized action. We shall now analyse this question in more detail.

We start with a minimally coupled massive scalar field. Equation (10.14) for the basis functions reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(a^3\frac{\mathrm{d}v_n}{\mathrm{d}t}\right)+a^3\left(\frac{n^2-1}{a^2}+m^2\right)v_n=0. \tag{10.24}$$

The appropriate solution to this equation is

$$v_n(t)=(\cosh Ht)^{-1}P^{-n}_{-\frac{1}{2}+\mathrm{i}\sqrt{m^2/H^2-9/4}}(\mathrm{i}\sinh Ht), \tag{10.25}$$

where P denotes an associated Legendre function of the first kind. The corresponding expression for (10.13) is given for a large mass m by

$$\Omega_n=a^2\left[\sqrt{n^2+m^2a^2}+\mathrm{i}\sinh Ht\left(1+\frac{1}{2}\frac{m^2a^2}{n^2+m^2a^2}\right)\right]+O\left(\frac{1}{m}\right). \tag{10.26}$$

The leading contribution to the amplitude of the decoherence factor is therefore

$$\ln|\boldsymbol{D}(t|\phi,\phi')|\simeq\frac{1}{4}\sum_{n=0}^{\infty}n^2\ln\frac{4a^2a'^2\sqrt{n^2+m^2a^2}\,\sqrt{n^2+m^2a'^2}}{\left(a^2\sqrt{n^2+m^2a^2}+a'^2\sqrt{n^2+m^2a'^2}\right)^2}. \tag{10.27}$$

The first term, n^2, in the sum comes from the degeneracy of the eigenfunctions. This expression has divergences that *cannot* be represented as additive functions of a and a'. This means that no one-argument counterterm to $\mathbf{\Gamma}$ and $\mathbf{\Gamma}'$ in (10.20) can cancel these divergences of the amplitude (Paz and Sinha 1992). One might try to apply standard regularization schemes from quantum field theory, such as dimensional regularization. The corresponding calculations have been performed in Barvinsky *et al.* (1999*a*) and will not be given here. The important result is that, although the sum (10.27) is found to be convergent, a *positive* value is obtained for it. This means that the decoherence factor must diverge for $(\phi-\phi')\to\infty$ and thus spoil one of the crucial properties of a density matrix—the boundedness of $\mathrm{tr}\,\hat{\rho}^2$. The dominant term in the decoherence factor reads

$$\ln|\boldsymbol{D}|=\frac{\pi}{24}(ma)^3+O(m^2),\ \ a\gg a', \tag{10.28}$$

and is thus unacceptable for a density matrix. Reduced density matrices are usually not considered in quantum field theory, so this problem has not been encountered before. A behaviour such as that in (10.28) is even obtained in the case of massless conformally invariant fields, for which one would expect a decoherence factor equal to one, since they decouple from the gravitational background. How, then, should one proceed in order to obtain a sensible regularization?

The crucial point is to perform a *redefinition* of the environmental fields and to invoke a physical principle to fix this redefinition. The situation is somewhat analogous to the treatment of the S-matrix in quantum field theory: the off-shell S-matrix and effective action depend on the parametrization of the quantum fields, in analogy to the non-diagonal elements of the reduced density matrix. Laflamme and Louko (1991) and Kiefer (1992) proposed, within special models, to rescale the environmental fields by a power of the scale factor. It was therefore suggested in Barvinsky *et al.* (1999*a*) to redefine the environmental fields by a power of the scale factor that corresponds to the conformal weight of the field (which is defined by the conformal invariance of the wave equation). For a scalar field in four space–time dimensions, this amounts to multiplication by a:

$$v_n(t) \to \tilde{v}_n(t) = a v_n(t), \tag{10.29}$$

$$\tilde{\Omega}_n = -\mathrm{i}a\frac{\mathrm{d}}{\mathrm{d}t}\ln \tilde{v}_n^*. \tag{10.30}$$

An immediate test of this proposal is to see whether the decoherence factor is equal to one for a massless conformally invariant field. In this case, the basis functions and frequency functions are, respectively,

$$\tilde{v}_n^*(t) = \left(\frac{1+\mathrm{i}\sinh Ht}{1-\mathrm{i}\sinh Ht}\right)^{n/2}, \tag{10.31}$$

$$\tilde{\Omega}_n = -\mathrm{i}a\frac{\mathrm{d}}{\mathrm{d}t}\ln \tilde{v}_n^*(t) = n. \tag{10.32}$$

Hence, $\tilde{\boldsymbol{D}}(t|\phi,\phi') \equiv 1$. The same holds for the electromagnetic field (which, in four space–time dimensions, is conformally invariant). It is interesting to note that the degree of decoherence caused by a certain field depends on the space–time dimension, since its conformal properties are dimension-dependent.

For a massive minimally coupled field, the new frequency function is

$$\tilde{\Omega}_n = \left[\sqrt{n^2+m^2a^2} + \mathrm{i}\sinh Ht\left(\frac{1}{2}\frac{m^2a^2}{n^2+m^2a^2}\right)\right] + O(1/m). \tag{10.33}$$

Note that, in contrast to (10.26), there is no factor of a^2 in front of this expression. Since (10.33) is valid in the large-mass limit, it corresponds to modes that evolve adiabatically on the gravitational background, the imaginary part in (10.33) describing particle creation.

It turns out that the imaginary part of the decoherence factor has at most logarithmic divergences and, therefore, affects only the phase of the density matrix. Moreover, these divergences decompose into an *additive* sum of one-argument functions and can thus be cancelled by adding counterterms to the classical action S_0 (and S_0') in (10.20) (Paz and Sinha 1992). The real part is simply convergent and gives a finite decoherence amplitude. This result is formally similar to the result for the decoherence factor in an application to quantum electrodynamics (Kiefer 1992).

For $a \gg a'$ (far off-diagonal terms), one gets the expression

$$|\tilde{\boldsymbol{D}}(t|\phi,\phi')| \simeq \exp\left[-\frac{(ma)^3}{24}\left(\pi - \frac{8}{3}\right) + O(m^2)\right]. \tag{10.34}$$

Compared with the naively regularized (and inconsistent) expression (10.28), π has effectively been replaced by $8/3 - \pi$. In the vicinity of the diagonal, one obtains

$$\ln|\tilde{\boldsymbol{D}}(t|\phi,\phi')| = -\frac{m^3\pi a(a-a')^2}{64}, \tag{10.35}$$

a behaviour similar to (10.34).

An interesting case is also provided by minimally coupled massless scalar fields, and gravitons. These share the following basis and frequency functions in their respective conformal parametrizations:

$$\tilde{v}_n^*(t) = \left(\frac{1+\mathrm{i}\sinh Ht}{1-\mathrm{i}\sinh Ht}\right)^{n/2}\left(\frac{n-\mathrm{i}\sinh Ht}{n+1}\right), \tag{10.36}$$

$$\tilde{\Omega}_n = \frac{n(n^2-1)}{n^2-1+H^2a^2} - \mathrm{i}\frac{H^2a^2\sqrt{H^2a^2-1}}{n^2-1+H^2a^2}. \tag{10.37}$$

They differ only in the range of the quantum number n ($2 \leq n$ for inhomogeneous scalar modes and $3 \leq n$ for gravitons) and in the degeneracies of the nth eigenvalue of the Laplacian,

$$\dim(n)_{\text{scal}} = n^2, \tag{10.38}$$

$$\dim(n)_{\text{grav}} = 2(n^2-4). \tag{10.39}$$

For far off-diagonal elements, one obtains the decoherence factor

$$|\tilde{\boldsymbol{D}}(t|\phi,\phi')| \sim \mathrm{e}^{-C(Ha)^3}, \quad a \gg a', \quad C > 0, \tag{10.40}$$

while in the vicinity of the diagonal one finds

$$|\tilde{\boldsymbol{D}}(t|\phi,\phi')| \sim \exp\left(-\frac{\pi^2}{32}(H-H')^2t^2\mathrm{e}^{4Ht}\right), \tag{10.41}$$

$$\sim \exp\left(-\frac{\pi^2H^4a^2}{8}(a-a')^2\right), \quad Ht \gg 1. \tag{10.42}$$

These expressions exhibit a rapid disappearance of non-diagonal elements during the inflationary evolution. The universe thus assumes classical properties at the onset of inflation. This justifies the use of classical cosmology since then.

The decohering influence of fermionic degrees of freedom has to be treated separately (Barvinsky *et al.* 1999*b*). It turns out that they are less efficient in producing decoherence. In the massless case, for example, their influence is completely absent.

The above analysis of decoherence was based on the state (10.11). One might, however, start with a quantum state which is a superposition of many semiclassical components, that is, many components of the form $\exp(\mathrm{i}S_0^k)$, where each S_0^k is a solution of the Hamilton–Jacobi equation for a and ϕ. Decoherence between different semiclassical *branches* of this kind has also been the subject of intense investigation (Halliwell 1989, Kiefer 1992). The important point is that decoherence between different branches is usually weaker than the above-discussed decoherence within one branch. Moreover, it usually follows from the presence of decoherence within one branch. In the special case of a superposition of (10.11) with its complex conjugate, one can immediately recognize that the decoherence between the semiclassical components is smaller than that within one component: in the expression (10.22) for the decoherence factor, the term $\Omega_n + \Omega_n'^*$ in the denominator is replaced by $\Omega_n + \Omega_n'$. Therefore, the imaginary parts of the frequency functions add up instead of partially cancelling each other, and

(10.22) becomes smaller. One also finds that the decoherence factor is equal to one for vanishing expansion of the semiclassical universe: in the context of a simple model, the decoherence factor between the $\exp(\mathrm{i}S_0)$ and $\exp(-\mathrm{i}S_0)$ components turns out to be given by (Kiefer 1992)

$$\exp\left(-\frac{\pi m H^2 a^3}{128\hbar}\right) \sim \exp\left(-10^{43}\right),$$

where m is the mass of a scalar field. The numerical value is obtained after some standard values for the parameters are inserted. Its smallness indicates that the present Universe can be treated classically to a high degree of accuracy.

We note that the decoherence between the $\exp(\mathrm{i}S_0)$ and $\exp(-\mathrm{i}S_0)$ components can be interpreted as a *symmetry breaking* in analogy to the case of sugar molecules (Joos *et al.* 2003). There, the Hamiltonian is invariant under space reflections, but the state of the sugar molecules exhibits chirality. Here, the Hamiltonian in the Wheeler–DeWitt equation is invariant under complex conjugation, while the 'actual states' (i.e. one decohering WKB component in the total superposition) are of the form $\exp(\mathrm{i}S_0)$ and are thus intrinsically complex. It is therefore not surprising that the recovery of the classical world follows only for complex states, in spite of the real nature of the Wheeler–DeWitt equation (see in this context Barbour 1993). Since this is a prerequisite for the derivation of the Schrödinger equation, one might even say that *time* (the WKB time parameter in the Schrödinger equation) arises from symmetry breaking.

The above considerations thus lead to the following picture. The universe was essentially 'quantum' at the onset of inflation. Mainly due to bosonic fields, decoherence set in and led to the emergence of many 'quasi-classical branches' which are dynamically independent of each other. Strictly speaking, the very concept of time makes sense only after decoherence has occurred. In addition to the horizon problem etc., inflation thus also solves the 'classicality problem'. It remains, of course, unclear why inflation happened in the first place (if it really did). If we were to look back from 'our universe' (our semiclassical branch) into the past, we would see, at the time of the onset of inflation, our component interfering with other components to form a timeless quantum-gravitational state. The universe would thus cease to be transparent to earlier times (because there was no time). This demonstrates in an impressive way that quantum-gravitational effects are not a priori restricted to the Planck scale.

The lesson to be drawn is thus that the universe can appear classical *only* if experienced from within. A hypothetical 'outside view' would see only a static quantum world. The most natural interpretation of quantum cosmology is an Everett-type interpretation, since the 'wave function of the universe' contains by definition all possible branches.[4] As macroscopic observers, however, we have access only to a tiny part of the cosmological wave function—the robust macroscopic branch which we follow. Incidentally, the original motivation for Everett to develop his interpretation was quantum gravity. To quote from Everett (1957, p. 454):

The task of quantizing general relativity raises serious questions about the meaning of the present formulation and interpretation of quantum mechanics when applied to so fundamental

[4]There also exist attempts to extend the 'Bohm interpretation' of quantum theory to quantum cosmology; cf. Pinto-Neto and Santini (2002) and Błaut and Kowalski-Glikman (1998).

a structure as the space-time geometry itself. This paper seeks to clarify the foundations of quantum mechanics. It presents a reformulation of quantum theory in a form believed suitable for application to general relativity.

This view is reflected by the words of Bryce DeWitt in DeWitt (1967*a*, p. 1141):

Everett's view of the world is a very natural one to adopt in the quantum theory of gravity, where one is accustomed to speak without embarrassment of the 'wave function of the universe.' It is possible that Everett's view is not only natural but essential.

10.2.3 Decoherence of primordial fluctuations

We have seen in the last subsection how important global degrees of freedom such as the scale factor of the universe can assume classical properties through interaction with irrelevant degrees of freedom such as density perturbations or gravitational waves. There are, however, situations when part of these 'irrelevant' variables become relevant themselves. According to the inflationary scenario of the early universe, all structure in the Universe arises from quantum fluctuations; see, for example, Liddle and Lyth (2000) and Börner (2003) for a review. If this is correct, we thus owe our existence entirely to the uncertainty relations. In order to serve as the seeds for structure (galaxies and the structures of galaxies), these quantum fluctuations have to become classical. Their imprint is seen in the anisotropy spectrum of the cosmic microwave background. The quantum-to-classical transition again relies heavily on the notion of decoherence; see Polarski and Starobinsky (1996) and Kiefer *et al.* (1998), and see Section 4.2.4 of Joos *et al.* (2003) for a review. Decoherence is effective when the wavelength of the primordial quantum fluctuations becomes much greater than the Hubble scale H_{I}^{-1} during the inflationary regime, where H_{I} denotes the Hubble parameter of inflation (which is approximately constant); cf. Fig. 7.6. The quantum state becomes strongly squeezed during this phase: the squeezing is in the field momentum, while the field amplitude becomes very broad. Such a state is highly sensitive to any interaction, even if small, with other ('environmental' or 'irrelevant') fields. It thereby decoheres into an ensemble of narrow wave packets that are approximately eigenstates of the field amplitude. A prerequisite is the classical nature of the background variables discussed in the last subsection, which is why one can talk about a 'hierarchy of classicality'. A detailed discussion of the classical pointer basis for the fluctuations is presented in Kiefer *et al.* (2007*a*).

Density fluctuations arise from the scalar part of the metric perturbations (plus the corresponding matter part). In addition, one has of course the tensor perturbations of the metric. These correspond to gravitons (Chapter 2). As with the scalar part, the tensor part evolves into a highly squeezed state during inflation, and decoherence happens for it too. The primordial gravitons should manifest themselves in a stochastic background of gravitational waves, which could probably be observed with the space-borne interferometer LISA to be launched in a few years' time. The observation of this background would constitute a direct test of linearized quantum gravity.

The decoherence time turns out to be of the order

$$t_{\mathrm{d}} \sim \frac{H_{\mathrm{I}}}{g}, \tag{10.43}$$

where g is a dimensionless coupling constant of the interaction with the other 'irrelevant' fields causing decoherence. The ensuing coarse-graining brought about by the decohering fields causes an entropy increase for the primordial fluctuations (Kiefer *et al.* 2000). The entropy production rate turns out to be given by $\dot{S} = H$, where H is the Hubble parameter of a general expansion. During inflation, H is approximately constant and the entropy increases linearly with t. In the post-inflationary phases (radiation- and matter-dominated universes), $H \propto t^{-1}$ and the entropy increases only logarithmically in time. The main part of the entropy of the fluctuations is thus created during inflation. Incidentally, this behaviour resembles the behaviour of chaotic systems, although no chaos is involved here. The role of the Lyapunov coefficient is played by the Hubble parameter, and the Kolmogorov entropy corresponds to the entropy production mentioned here.

Decoherence also plays an important role for quantum black holes and in the context of wormholes and string theory; see Section 4.2.5 of Joos *et al.* (2003). The idea of understanding the small observed positive value of the cosmological constant on the basis of decoherence is presented in Kiefer *et al.* (2011).

10.3 Arrow of time

One of the most intriguing open problems is the origin of irreversibility in our Universe, also called the problem of the arrow of time. Since quantum gravity may provide the key to its solution, this topic will be briefly reviewed here. More details and references can be found in Zeh (2007).

Although most of the fundamental laws of nature do not distinguish between past and future, there are many classes of phenomena which exhibit an arrow of time. This means that their time-reversed version is, under ordinary conditions, never observed. The most important arrows of time are the following:

- The radiation arrow (advanced versus retarded radiation).
- The Second Law of Thermodynamics (increase of entropy).
- Quantum theory (the measurement process and emergence of classical properties).
- Gravitational phenomena (expansion of the Universe and emergence of structure by gravitational condensation).

All these arrows point in the same direction. The expansion of the Universe is distinguished because it does not refer to a class of phenomena; it is a single process. It has therefore been suggested that it is the common root of all other arrows of time—the 'master arrow'. We shall see in the course of our discussion that this seems indeed to be the case. But first we shall consider the various arrows of time in more detail.

The *radiation arrow* is characterized by the fact that fields interacting with local sources are usually described by *retarded* solutions, which in general lead to a damping of the source. Advanced solutions are excluded. These would describe the reversed process, during which the field propagates coherently towards its source, leading to its excitation instead of damping. This holds, in fact, for all wave phenomena. In electrodynamics, a solution of Maxwell's equations can be described by

$$\begin{aligned} A^\mu &= \text{source term plus boundary term} \\ &= A^\mu_{\text{ret}} + A^\mu_{\text{in}} \\ &= A^\mu_{\text{adv}} + A^\mu_{\text{out}}, \end{aligned}$$

where A^μ is the vector potential. The important question is, then, why the observed phenomena obey $A^\mu \approx A^\mu_{\text{ret}}$ or, in other words, why

$$A^\mu_{\text{in}} \approx 0 \tag{10.44}$$

holds instead of $A^\mu_{\text{out}} \approx 0$. Equation (10.44) is called a 'Sommerfeld radiation condition'. It is believed that the radiation arrow can be traced back to thermodynamics: due to the absorption properties of the material which constitutes the walls of the laboratory in which electrodynamic experiments are performed, ingoing fields will be absorbed within a very short time and (10.44) will be fulfilled. The Second Law of Thermodynamics (see below) is responsible for the thermal properties of absorbers.

The condition (10.44) also seems to hold for the Universe as a whole (the 'darkness of the night sky'). This so-called Olbers' paradox (why the sky at night is dark) can be solved by noting that the Universe is, in fact, not static, but has a finite age and is much too young to have enough stars for a bright night sky. This is, of course, not sufficient to understand the validity of (10.44) for the Universe as a whole. At an early stage, the Universe was a hot plasma in thermal equilibrium. Only the expansion of the Universe and the ensuing redshift of the radiation are responsible for the fact that radiation has decoupled from matter and cooled to its present value of about three kelvin—the temperature of the approximately isotropic cosmic background radiation with which the night sky 'glows'. During the expansion, a strong thermal non-equilibrium could develop, which enabled the formation of structure.

The second arrow is described by the *Second Law of Thermodynamics*: for a closed system, entropy does not decrease. The total change of entropy is given by

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \underbrace{\left(\frac{\mathrm{d}S}{\mathrm{d}t}\right)_{\text{ext}}}_{\mathrm{d}S_{\text{ext}}=\delta Q/T} + \underbrace{\left(\frac{\mathrm{d}S}{\mathrm{d}t}\right)_{\text{int}}}_{\geq 0},$$

so that according to the Second Law the second term is non-negative. As the increase of entropy is also relevant to physiological processes, the Second Law is responsible for the subjective experience of irreversibility, in particular the ageing process. If applied to the Universe as a whole, it would predict an increase of its total entropy, which would seem to lead to its 'heat death' (*Wärmetod*).

The laws of thermodynamics are based on microscopic statistical laws that are time-symmetric. How can the Second Law be derived from such laws? As early as in the nineteenth century, objections were formulated against a statistical foundation of the Second Law. These were, in particular,

- Loschmidt's reversibility objection (*Umkehreinwand*), and
- Zermelo's recurrence objection (*Wiederkehreinwand*).

Loschmidt's objection states that a reversible dynamics must lead to an equal number of transitions from an improbable to a probable state and from a probable to an improbable state. With overwhelming probability, the system should be in its most probable state, that is, in thermal equilibrium. Zermelo's objection is based on a theorem by Poincaré, according to which every system comes arbitrarily close to its initial state (and therefore to its initial entropy) after a finite amount of time. This objection is irrelevant, since the corresponding 'Poincaré times' are already greater than the age of the Universe for systems with only a few particles. The reversibility objection can only be avoided if a special boundary condition of low entropy holds for the early Universe. Therefore, for the derivation of the Second Law, one needs a special *boundary condition*.

Such a boundary condition must either be postulated or be derived from a fundamental theory. The formal description of entropy increase from such a boundary condition is done by master equations; cf. Joos *et al.* (2003). These are equations for the 'relevant' (coarse-grained) part of the system. In an open system, the entropy can of course decrease, provided the entropy capacity of the environment is large enough to at least compensate this entropy decrease. This is crucial for the existence of life, and one particularly efficient process in this respect is photosynthesis. The huge entropy capacity of the environment comes in this case from the high temperature gradient between the hot Sun and the cold, empty space: few high-energy photons (with low entropy) arrive on the Earth, while many low-energy photons (with high entropy) leave it. Therefore, the thermodynamic arrow of time also points towards cosmology: how can gravitationally condensed objects such as the Sun arise in the first place?

A third important arrow of time is the *quantum-mechanical arrow*. The Schrödinger equation is time-reversal invariant, but the measurement process, either through

- a dynamical *collapse* of the wave function or
- an Everett *branching*,

distinguishes a direction; cf. Section 10.2. We have seen that growing entanglement with other degrees of freedom leads to decoherence. The local entropy thereby increases. Again, decoherence only works if a special initial condition—a condition of weak entanglement—holds. But where can this come from?

The last of the main arrows is the *gravitational arrow* of time. Although the Einstein field equations are time-reversal invariant, gravitational systems in nature distinguish a certain direction: the Universe as a whole *expands*, while local systems such as stars form by *contraction*, for example from gas clouds. It is by this gravitational contraction that the high temperature gradients between stars such as the Sun and the empty space arise. Because of the negative heat capacity of gravitational systems, homogeneous states possess a low entropy, whereas inhomogeneous states possess a high entropy—just the opposite of non-gravitational systems.

But where does the initial homogeneous state needed for the gravitational collapse come from? In the nineteenth century, Ludwig Boltzmann speculated that the Second Law had its origin in a gigantic fluctuation in the Universe. His picture was that the Universe exists eternally and has maximal entropy most of the time, but that on very rare occasions (which, of course, can happen in an eternal Universe) the entropy fluctuates to a very low value, from which it will then increase; this would then enable our existence and would lead to the arrow of time that we observe. The weak point in this argument was pointed out by von Weizsäcker (1939): if one took the possibility of entropy fluctuations into account, a fluctuation that produced at once the world that we observe, including our existence and our memories, although by itself extremely unlikely, would still be much more probable than Boltzmann's fluctuation, which has to create the whole history of the world in addition to the present state. But von Weizsäcker's argument can also be criticized: if we live in a gigantic (maybe infinite) universe (also called a 'multiverse', see below), it would be much more probable that isolated brains popped occasionally into existence, which would not see a structured and consistent world as we do, but a world ruled by chaos. Such isolated brains are called 'Boltzmann brains'; see De Simone *et al.* (2010) and the references therein. Since we are not Boltzmann brains, it makes sense to assume that we have really evolved from a low-entropy initial state at the big bang. Our Universe thus has started in a highly special way—but how special?

To answer this question, one has to calculate the actual entropy of our Universe and its maximum possible entropy. The non-gravitational entropy is dominated by the photons of the cosmic microwave background radiation; it contributes about $2 \times 10^{89} k_{\mathrm{B}}$ (Egan and Lineweaver 2010). Linde and Vanchurin (2010) have, moreover, given an estimate of an upper limit for the non-gravitational entropy, which would be obtained if all particles were ultra-relativistic: their value is about $10^{90} k_{\mathrm{B}}$ and thus only about one order of magnitude higher than the CMB value. These are very large numbers, but they are much smaller than the gravitational contribution to the entropy. Unfortunately, a general expression for gravitational entropy does not exist. An exception is the Bekenstein–Hawking entropy (7.23), which describes the most extreme form of gravitational collapse: the collapse to a black hole.

To see how large the Bekenstein–Hawking entropy can become, let us estimate its value for the Galactic Black Hole—the supermassive black hole in the centre of our Milky Way, with a mass $M \approx 3.9 \times 10^6 M_{\odot}$. Neclecting its angular momentum, which would anyway decrease the estimated entropy, one gets from (7.23), setting $k_{\mathrm{B}} = 1$ from now on,

$$S_{\mathrm{GBH}} = \pi \left(\frac{R_{\mathrm{S}}}{l_{\mathrm{P}}} \right)^2 \approx 6.7 \times 10^{90}, \tag{10.45}$$

where R_{S} denotes the Schwarzschild radius. This already exceeds by more than one order of magnitude the non-gravitational contribution to the entropy of the observable Universe. According to Egan and Lineweaver (2010), all supermassive black holes together yield an entropy of $S = 3.1^{+3.0}_{-1.7} \times 10^{104}$.

Roger Penrose has pointed out that the maximum entropy for the observable Universe would be obtained if all its matter were assembled into one black hole (Penrose 1981). Taking the most recent observational data, this would yield the entropy (the calculation can be found in the appendix of Kiefer (2009))

$$S_{\mathrm{max}} \approx 1.8 \times 10^{121}. \tag{10.46}$$

This may not yet be the maximum possible entropy. Our Universe currently exhibits an acceleration which could be caused by a cosmological constant Λ. If this were true, it would expand forever, and the entropy in the far future would be dominated by the entropy of the cosmological event horizon. This entropy is called the 'Gibbons–Hawking entropy' (Gibbons and Hawking 1977). The calculation can be found in the appendix of Kiefer (2009), and the result is

$$S_{\mathrm{GH}} = \frac{3\pi}{\Lambda l_{\mathrm{P}}^2} \approx 2.9 \times 10^{122}, \tag{10.47}$$

which is about one order of magnitude higher than (10.46).

Following the arguments in Penrose (1981), the 'probability' of our Universe can be estimated as

$$\frac{\exp(S)}{\exp(S_{\max})} \approx \frac{\exp(3.1 \times 10^{104})}{\exp(2.9 \times 10^{122})} \approx \exp(-2.9 \times 10^{122}). \tag{10.48}$$

Our Universe is thus very special indeed. It must have 'started' near the big bang with an extremely low entropy; the Universe must have been very smooth in the past, with no white holes being present. Penrose has reformulated this observation in his Weyl-tensor hypothesis: the Weyl tensor is zero near the big bang (describing a smooth state), but diverges in a big crunch (provided the Universe recollapses). Since the Weyl tensor describes, in particular, gravitational waves, this hypothesis entails that all gravitational waves must be retarded. This is analogous to the Sommerfeld condition in electrodynamics and the absence of advanced electromagnetic waves (Zeh 2007). There, the electromagnetic arrow can be traced back to the thermodynamic arrow and the Second Law by using the thermodynamic properties of absorbers, but this is not possible here, because the absorption cross-section for gravitational waves is too small.

The Weyl-tensor hypothesis is not yet an explanation, but only a description of the low initial cosmic entropy. We shall adopt the point of view that a fundamental theory of quantum gravity may provide the desired explanation. The formalism of quantum cosmology (Chapter 8) is especially suited for this purpose. Let us see whether the origin of irreversibility can be traced back to the structure of the Wheeler–DeWitt equation (or its generalization to a difference equation in loop quantum cosmology). Taking as in Section 8.2 a minisuperspace model with additional small inhomogeneities in the form of a multipole expansion, we arrive at a Wheeler–DeWitt equation of the form

$$\hat{H}\,\Psi = \left(\frac{2\pi G\hbar^2}{3}\frac{\partial^2}{\partial\alpha^2} + \sum_i \left[-\frac{\hbar^2}{2}\frac{\partial^2}{\partial x_i^2} + \underbrace{V_i(\alpha, x_i)}_{\to 0 \text{ for } \alpha \to -\infty} \right] \right) \Psi = 0, \tag{10.49}$$

where the $\{x_i\}$ denote the inhomogeneous degrees of freedom and the neglected homogeneous ones; $V_i(\alpha, x_i)$ are the corresponding potentials. One recognizes immediately that this Wheeler–DeWitt equation is hyperbolic with respect to the intrinsic time α. Initial conditions are thus most naturally formulated with respect to constant α.

The important property of (10.49) is that the potential becomes small for $\alpha \to -\infty$ (where the classical singularities would occur), but complicated for increasing α. In the general case (not restricting ourselves to small inhomogeneities), a further motivation for this assertion is the BKL conjecture, according to which spatial gradients become small near a space-like singularity. The Wheeler–DeWitt equation thus possesses an asymmetry with respect to the 'intrinsic time' α. One can in particular impose the simple boundary condition (Zeh 2007)

$$\Psi \quad \overset{\alpha \to -\infty}{\longrightarrow} \quad \psi_0(\alpha) \prod_i \psi_i(x_i), \tag{10.50}$$

which means that the degrees of freedom are initially *not* entangled. If (10.50) were taken as an 'initial condition', the Wheeler–DeWitt equation would, through the occurrence of the potentials in (10.49), lead to a wave function that for increasing α becomes *entangled* between α and all modes. This, then, would lead to an increase of the local entanglement entropy, that is, an increase of the entropy which is connected with the subset of 'relevant' degrees of freedom. Calling the latter $\{y_i\}$, one has

$$S(\alpha, \{y_i\}) = -k_{\rm B}\, {\rm tr}(\rho \ln \rho), \tag{10.51}$$

where ρ is the reduced density matrix corresponding to α and $\{y_i\}$. This is obtained by tracing out all irrelevant degrees of freedom in the full wave function. Because ρ becomes more mixed if the entanglement is stronger, entropy would increase with increasing scale factor. Since there is a close connection between entanglement entropy and thermodynamical entropy (Peres 1995, Chapter 9), this could be the origin of the gravitational arrow of time. In the semiclassical limit, where the time parameter t is constructed from α (and other degrees of freedom) (cf. Section 5.4), entropy would increase with increasing t. This, then, would *define* the direction of time and would be the origin of the observed irreversibility in the world. The expansion of the Universe would then be a tautology. Due to the increasing entanglement, the Universe would rapidly assume classical properties for the relevant degrees of freedom due to decoherence. The gravitational arrow of time would then really be the root of both the quantum-mechanical and the thermodynamical arrow of time. Quantum gravity could thus provide the master arrow, the formal reason being the asymmetric appearance of α in the Wheeler–DeWitt equation. An analysis of this kind should, however, not be restricted to quantum geometrodynamics, but should also be applicable to loop quantum gravity and to string theory.

At the present level of understanding, the initial condition (10.50) is only one possibility out of many others; no principle is yet known that could enforce it. It would be an intriguing idea if a full mathematical understanding of canonical quantum gravity allowed only one solution (or a small class of solutions) that would automatically fulfill (10.50). If this were the case, one would have derived the arrow of time from quantum gravity.

In the case of a classically recollapsing universe, the boundary condition (10.50) has interesting consequences: since it is formulated at $\alpha \to \infty$, increasing entropy is always correlated with increasing α, that is, increasing size of the universe; cf. also Fig. 8.1. Consequently, the arrow of time formally reverses near the classical turning point (Kiefer and Zeh 1995). It turns out that this region is then fully quantum, so no paradox arises; it just means that there are many quasi-classical components of the wave function, each describing a universe that is experienced from within as expanding.

All these components interfere destructively near the classical turning point, which then constitutes the 'end' of evolution. This is analogous to the quantum region at the onset of inflation discussed in Section 10.2.2.

The above ideas may, with slight elaborations, also apply to the idea of a *multiverse*. A multiverse is a gigantic universe (much larger than the observable part of our Universe) with many approximately homogeneous sub-universes; see Linde and Vanchurin (2010) and the references therein. Quantum entanglement is not limited to sub-horizon scales and may thus be effective also in the full multiverse. Decoherence should then lead to the same arrow of time everywhere in the multiverse. By applying to the multiverse the idea that quantum fluctuations, after their effective classicality due to decoherence, become the seeds for galaxy formation, Linde and Vanchurin (2010) estimated the number of realizations of the emergent classical fluctuations in the gigantic multiverse. This number would also correspond to the number of branches of the universal wave function in the Everett interpretation when applied to our Hubble domain. After decoherence, each realization can serve as a classical initial condition for the subsequent evolution of the universe. Linde and Vanchurin found for the total number of locally distinguishable 'Friedmann universes' the estimate

$$\mathrm{e}^{S_{\mathrm{pert}}} \lesssim \mathrm{e}^{\mathrm{e}^{3N}}, \tag{10.52}$$

where S_{pert} is the total entropy of the perturbations, and N is the number of e-folds of slow-roll (post-eternal) inflation. In the simplest models of chaotic inflation, one gets the incredibly high number

$$10^{10^{10^7}}.$$

(A much lower number—with two instead of three exponentials—is obtained in the case of a positive cosmological constant.) If one adopts, in addition, the landscape picture of string theory, this estimate would correspond to the case of one vacuum. If one took all the vacua into account, the number would be even higher. The issue of the Wheeler–DeWitt equation on a configuration space mimicking the landscape picture is discussed, for example, in Bouhmadi-López and Moniz (2007) and the references therein.

10.4 Reflections and outlook

'There is no experimental evidence for the quantization of the gravitational field, but we believe quantization should apply to all the fields of physics. They all interact with one another, and it is difficult to see how some could be quantized and others not.' This quote from Dirac (1968) still reflects the prevailing opinion. In fact, the main problem so far is the lack of experimental tests that could guide us towards the correct theory. Dyson (2004) argued that it may be impossible to observe the existence of individual gravitons; cf. von Borzeszkowski and Treder (1988) for a similar standpoint. Even if this were true (which is far from clear), the experimental testability of quantum gravity would go much beyond the observation of single gravitons. We have encountered in this book various concrete predictions, such as corrections to Newton's law of gravitational attraction and effects in the CMB anisotropy spectrum. We have also encountered various ideas of quantum-gravity phenomenology. There, one looks for possible violations of established laws, independent of the origin of such violations; they could arise from a new purely classical extension of GR, but their root could also lie in quantum gravity.

We are thus optimistic that it will eventually be possible to test a theory of quantum gravity. As long as this is not possible, one has to focus on conceptual and mathematical problems as concretely as possible; to quote from Dyson (2004), 'It is better to be wrong than to be vague.' We have seen in the preceding chapters that the main problem, both conceptually and mathematically, is the background independence of quantum gravity: we cannot take fields on a rigid, non-dynamical background space–time as usual and quantize them, but we have to devise a full quantum theory that *includes* the 'background', that is, the gravitational degrees of freedom.

This lack of background supports a fundamental *relational* point of view (Barbour 2000, Rovelli 2004). In the philosophy of science, this can be embedded into a direction called structural realism. To quote from Rickles and French (2006, p. 25):

> Structuralism shifts the focus onto the relational structures themselves and away from the objects, which must then be reconceived, in some sense, from the structure.

In Section 1.3, we have referred to Isham's distinction into primary and secondary theories of quantum gravity. Primary theories are theories that result from a direct 'quantization' of GR. The classical analyis of GR has already disclosed its dynamical degrees of freedom: these are the three-dimensional geometry, and the configuration space is superspace. Whereas in the classical theory three-dimensional spaces can be stacked together to give a space–time, this is no longer possible in the quantum theory. The wave functional is defined on superspace (the space of all three-geometries), and there is no trace of space–time at the most fundamental level. Since space is still present in the form of a dynamical three-geometry, time has disappeared. This disappearance of time goes beyond pure background independence. The situation is similar to that in mechanics, where classically trajectories exist (the analogue of space–times), but where quantum-mechanically only the configuration (the analogue of three-geometry) remains. The disappearance of time is thus a very conservative conclusion, which is drawn from a straightforward application of quantization recipes to GR. To quote from Wheeler (1968, p. 253):

> These considerations reveal that the concepts of spacetime and time itself are not primary but secondary ideas in the structure of physical theory. These concepts are valid in the classical approximation. However, they have neither meaning nor application under circumstances when quantum-geometrodynamical effects become important. ... There is no spacetime, there is no time, there is no before, there is no after. The question what happens 'next' is without meaning.

A similar statement can be found in DeWitt (1999, p. 7):

> ...one learns that time and probability are both *phenomenological* concepts.

The reference to probability refers to the 'Hilbert-space problem', which is intimately connected with the 'problem of time'. If time is absent, the notion of a probability conserved in time does not make much sense; the traditional Hilbert-space structure was designed to implement the probability interpretation, and its fate in a timeless world thus remains open.

This conceptual discussion is not restricted to the Wheeler–DeWitt equation. The same applies to loop quantum gravity, which is a variant of the canonical approach and is thus again centred around constraint equations. Most likely, these problems are also present in string theory, although there the situation is less clear. What is clear is that string theory contains GR in an appropriate limit and should thus also contain its quantized form, the Wheeler–DeWitt equation, in some limit; string theory should thus face the problem of time too. However, string theory goes far beyond this. Due to its apparent implementation of the holographic principle, it also provides new insights into space. In a sense, space (or some part of it) vanishes too. Whether there is some connection to the effective one-dimensionality of space at small scales in some approaches to quantum GR (asymptotic safety, dynamical triangulation, etc.) is open.

Einstein, in his later years, emphasized the important role of the connection in GR, which he considered to be superior to the metric. About two weeks before his death, he wrote the following sentences (Einstein 1955):[5]

> ...the essential achievement of general relativity, namely to overcome 'rigid' space (i.e. the inertial frame), is *only indirectly* connected with the introduction of a Riemannian metric. The directly relevant conceptual element is the 'displacement field' (Γ^l_{ik}), which expresses the infinitesimal displacement of vectors. It is this which replaces the parallelism of spatially arbitrarily separated vectors fixed by the inertial frame (i.e. the equality of corresponding components) by an infinitesimal operation. This makes it possible to construct tensors by differentiation and hence to dispense with the introduction of 'rigid' space (the inertial frame). In the face of this, it seems to be of secondary importance in some sense that some particular Γ field can be deduced from a Riemannian metric ....

To overcome 'rigid' space is, of course, nothing else than to search for a background-independent formulation.

In spite of this higher status of Γ compared with the metric, the arena for the quantum geometrodynamical wave function is the three-dimensional *metric*, not the *connection*. It thus seems that, at least in the most straightforward quantization scheme, the metric is more important in the quantum theory. The situation is somewhat different in loop quantum gravity. There, the central configuration variable is the holonomy built from the connection (4.121), which is a linear combination of the three-dimensional connection and the extrinsic curvature. In this sense, loop quantum gravity is perhaps closer to Einstein's ideas. The discrete structure of space as predicted by it can be traced back to this special choice of variables. One important consequence is that loop quantum gravity seems to be more promising with respect to the avoidance of classical singularities.

There is little evidence so far that the direct quantization of GR leads to a prediction of low-energy coupling constants and masses. The original hope was that string

[5] '...die wesentliche Leistung der allgemeinen Relativitätstheorie, nämlich die Überwindung (des) "starren" Raumes, d.h. des Inertialsystems, ist *nur indirekt* mit der Einführung einer Riemann-Metrik verbunden. Das unmittelbar wesentliche begriffliche Element ist das die infinitesimale Verschiebung von Vektoren ausdrückende "Verschiebungsfeld" (Γ^l_{ik}). Dieses nämlich ersetzt den durch das Inertialsystem gesetzten Parallelismus räumlich beliebig getrennter Vektoren (nämlich Gleichheit entsprechender Komponenten) durch eine infinitesimale Operation. Dadurch wird die Bildung von Tensoren durch Differentiation ermöglicht und so die Einführung des "starren" Raumes (Inertialsystem) entbehrlich gemacht. Dem gegenüber erscheint es in gewissem Sinne von sekundärer Wichtigkeit, dass ein besonderes Γ-Feld sich aus der Existenz einer Riemann-Metrik deduzieren lässt ....'

theory could achieve this goal, because it contains only one fundamental dimensionful parameter, the string length. But as is well known, only a fine-tuned combination of the low-energy constants leads to a universe like ours in which human beings can exist. It would thus appear strange if a fundamental theory possessed just the right constants to achieve this. Hogan (2000) argued that grand unified theories constrain relations among parameters, but leave enough freedom for a selection. In particular, he suggested that one coupling constant and two light fermion masses are *not* fixed by the symmetries of the fundamental theory. One could then determine these remaining free constants only by the (weak form of the) *anthropic principle*: they have values such that a universe like ours is possible. The cosmological constant, for example, must not be much greater than the presently observed value, because otherwise the universe would expand much too fast to allow the formation of galaxies. The Universe is, however, too special to be explainable on purely anthropic grounds. From the anthropic principle alone, one would not need such a special universe. In the case of the cosmological constant, one could imagine its calculation from a fundamental theory. Taking the presently observed value for Λ, one can construct a mass according to

$$\left(\frac{\hbar^2\Lambda^{1/2}}{G}\right)^{1/3} \approx 15 \text{ MeV}, \tag{10.53}$$

which in elementary particle physics is not an unusally large or small value. The observed value of Λ could thus emerge together with medium-size particle mass scales. Weinberg (2007) has instead suggested that the cosmological constant and the masses that set the sale of electroweak symmetry breaking can only be determined anthropically, while the remaining masses and constants can then be calculated from a fundamental theory.

Since fundamental theories are expected to contain only one dimensionful parameter, low-energy constants emerge, in fact, from fundamental quantum *fields*. An important example in string theory is the dilaton field, from which one can in principle calculate the gravitational constant. In order for these fields to mimic physical constants, two conditions have to be satisfied. First, decoherence must be effective in order to guarantee a classical behaviour of the field. Second, this 'classical' field must then be approximately constant in large enough space–time regions, within the limits given by experimental data. The field may still vary over large times or large spatial regions and thus mimic a 'time- or space-varying constant'; cf. Uzan (2003).

Some physicists entertain the idea that gravity is not represented by a fundamental dynamical field, but is only an effective theory; see, for example, Padmanabhan (2010) and the references therein. The main motivation for this speculation is the presence of thermodynamical analogies for space–time horizons (Chapter 7). The situation would then be analogous to that in hydrodynamics, which is governed by effective equations for density, pressure, fluid velocity, and so on, and not by fundamental equations referring to microscopic variables. One therefore does not attempt to quantize the effective equations. If the gravitational degrees of freedom were only effective too, the same objection would apply and practically all of the approaches discussed in this book would become obsolete. While this possibility cannot be excluded, it seems highly unlikely to me. The analogy of gravitational waves with electromagnetic waves in the weak-field limit is so strong that it is hard to see why one is of quantum nature and not the other; why photons exist and gravitons do not. Moreover, GR does not resemble hydrodynamics, because in the Einstein equations the metric occurs at the same dynamical level as the fundamental fields in the energy–momentum tensor.

In this book, I have presented two main approaches to quantum gravity—quantum GR (in both covariant and canonical versions) and string theory. Both rely on the linear structure of quantum theory, that is, the general validity of the superposition principle. In this sense, string theory too is a rather conservative approach, in spite of its 'exotic' features such as higher dimensions. This has to be contrasted with the belief of some of the founders of quantum mechanics (especially Heisenberg) that quantum theory would break down even on going from the level of atoms to the level of nuclei. It is, of course, imaginable that a fundamental theory of quantum gravity is intrinsically non-linear (Penrose 1996, Singh 2005). This is in contrast to most currently studied theories of quantum gravity (the ones discussed in this book), which are linear. However, so far no empirical evidence for any non-linear modification of quantum theory exists.

Quantum gravity has been studied since the end of the 1920s. Undoubtedly, much progress has been made since then. I hope that this book has given some impressions of this progress. The final goal has not yet been reached. The belief expressed here is that a consistent and experimentally successful theory of quantum gravity will be available in the future. However, it may still take a while before that time is reached.

Further reading: Amelino-Camelia and Kowalski-Glikman (2005), Carr (2007), Joos *et al.* (2003), Zeh (2007).

References

Abdo, A. A. and 207 other authors (2009). A limit on the variation of the speed of light arising from quantum gravity effects. *Nature*, **462**, 331–4.

Achúcarro, A. and Townsend, P. K. (1986). A Chern-Simons action for three-dimensional anti-de Sitter supergravity theories. *Phys. Lett. B*, **180**, 89–92.

Agullo, I., Barbero G., J. F., Borja, E. F., Diaz-Polo, J., and Villaseñor, E. J. S. (2010). Detailed black hole state counting in loop quantum gravity. *Phys. Rev. D*, **82**, 084029 [31 pp.].

Aharony, O., Gubser, S. S., Maldacena, J., Ooguri, H., and Oz, Y. (2000). Large N field theories, string theory and gravity. *Phys. Rep.*, **323**, 183–386.

Ahmed, M., Dodelson, S., Greene, P. B., and Sorkin, R. (2004). Everpresent Λ. *Phys. Rev. D*, **69**, 103523 [8 pp.].

Albers, M., Kiefer, C., and Reginatto, M. (2008). Measurement analysis and quantum gravity. *Phys. Rev. D*, **78**, 064051 [17 pp.].

Al'tshuler, B. L. and Barvinsky, A. O. (1996). Quantum cosmology and physics of transitions with a change of the spacetime signature. *Phys. Usp.*, **39**, 429–59.

Ambjørn, J. and Loll, R. (1998). Non-perturbative Lorentzian quantum gravity, causality and topology change. *Nucl. Phys. B*, **536**, 407–34.

Ambjørn, J., Jurkiewicz, J., and Loll, R. (2005). Reconstructing the Universe. *Phys. Rev. D*, **72**, 064014 [24 pp.].

Ambjørn, J., Jurkiewicz, J., and Loll, R. (2010). In *New paths towards quantum gravity* (ed. B. Booß-Bavnbek, G. Esposito, and M. Lesch), pp. 59–124. Lecture Notes in Physics 807. Springer, Berlin.

Amelino-Camelia, G. and Kowalski-Glikman, J. (eds.) (2005). *Planck scale effects in astrophysics and cosmology.* Lecture Notes in Physics 669. Springer, Berlin.

Amelino-Camelia, G. and Stachel, J. (2009). Measurement of the space–time interval between two events using the retarded and advanced times of each event with respect to a time-like world line. *Gen. Relativ. Gravit.*, **41**, 1107–24.

Amelino-Camelia, G., Ellis, J., Mavromatos, N. E., Nanopoulos, D. V., and Sarkar, S. (1998). Tests of quantum gravity from observations of γ-ray bursts. *Nature*, **393**, 763–5.

Amelino-Camelia, G., Lämmerzahl, C., Macias, A., and Müller, H. (2005). The search for quantum gravity signals. *AIP Conf. Proc.*, **758**, 30–80.

Anderson, J. (1967). *Principles of relativity physics.* Academic Press, New York.

Anderson, E. (2007*a*). On the recovery of geometrodynamics from two different sets of first principles. *Stud. Hist. Phil. Mod. Phys.*, **38**, 15–57.

Anderson, E. (2007*b*). Emergent semiclassical time in quantum gravity. I. Mechanical models. *Class. Quantum Grav.*, **24**, 2935–78.

Anderson, E. (2007*c*). Emergent semiclassical time in quantum gravity. II. Full geometrodynamics and minisuperspace examples. *Class. Quantum Grav.*, **24**, 2979–3004.

Anselmi, D. (2003). Absence of higher derivatives in the renormalization of propagators in quantum field theories with infinitely many couplings. *Class. Quantum Grav.*, **20**, 2355–78.

Antoniadis, I., Arkani-Hamed, N., Dimopoulos, S., and Dvali, G. (1998). New dimensions at a millimeter to a fermi and superstring at TeV. *Phys. Lett. B*, **436**, 257–63.

Arkani-Hamed, N., Dimopoulos, S., and Dvali, G. (1998). The hierarchy problem and new dimensions at a millimeter. *Phys. Lett. B*, **429**, 263–72.

Arnowitt, R., Deser, S., and Misner, C. W. (1962). The dynamics of general relativity. In *Gravitation: an introduction to current research* (ed. L. Witten), pp. 227–65. Wiley, New York.

Ashtekar, A. (1986). New variables for classical and quantum gravity. *Phys. Rev. Lett.*, **57**, 2244–7.

Ashtekar, A. (1988). *New perspectives in canonical gravity.* Bibliopolis, Naples.

Ashtekar, A. (1991). *Lectures on non-perturbative canonical gravity.* World Scientific, Singapore.

Ashtekar, A. and Geroch, R. (1974). Quantum theory of gravitation. *Rep. Prog. Phys.*, **37**, 1211–56.

Ashtekar, A. and Lewandowski, J. (1997). Quantum theory of geometry I: area operators. *Class. Quantum Grav.*, **14**, A55–82.

Ashtekar, A. and Lewandowski, J. (1998). Quantum theory of geometry II: volume operators. *Adv. Theor. Math. Phys.*, **1**, 388–429.

Ashtekar, A. and Lewandowski, J. (2004). Background independent quantum gravity: a status report. *Class. Quantum Grav.*, **21**, R53–152.

Ashtekar, A. and Pierri, M. (1996). Probing quantum gravity through exactly soluble midi-superspaces I. *J. Math. Phys.*, **37**, 6250–70.

Ashtekar, A. and Singh, P. (2011). Loop quantum cosmology: a status report. arXiv: 1108.0893v1 [gr-qc] [136 pp.].

Ashtekar, A., Marolf, D., and Mourão, J. (1994). Integration on the space of connections modulo gauge transformations. In *Proc. Cornelius Lanczos Int. Centenary Conf.* (ed. J. D. Brown *et al.*), pp. 143–60. SIAM, Philadelphia.

Ashtekar, A., Baez J., Corichi, A., and Krasnov, K. (1998). Quantum geometry and black hole entropy. *Phys. Rev. Lett.*, **80**, 904–7.

Ashtekar, A., Krasnov, K., and Baez, J. (2000). Quantum geometry of isolated horizons and black hole entropy. *Adv. Theor. Math. Phys.*, **4**, 1–94.

Ashtekar, A., Bojowald, M., and Lewandowski, L. (2003). Mathematical structure of loop quantum cosmology. *Adv. Theor. Math. Phys.*, **7**, 233–68.

ATLAS Collaboration (2011). Search for new physics in dijet mass and angular distributions in *pp* collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector. arXiv:1103.3864v1 [51 pp.].

Audretsch, J., Hehl, F. W., and Lämmerzahl, C. (1992). Matter wave interferometry and why quantum objects are fundamental for establishing a gravitational theory.

In *Relativistic gravity research* (ed. J. Ehlers and G. Schäfer), pp. 368–407. Lecture Notes in Physics 410. Springer, Berlin.

Babichev, E., Deffayet, C., and Ziour, R. (2010). Recovery of general relativity in massive gravity via the Vainshtein mechanism. *Phys. Rev. D*, **82**, 104008 [27 pp.].

Baierlein, R. F., Sharp, D. H., and Wheeler. J. A. (1962). Three-dimensional geometry as carrier of information about time. *Phys. Rev.*, **126**, 1864–65.

Baekler, P. and Hehl, F. W. (2011). Beyond Einstein–Cartan gravity: quadratic torsion and curvature invariants with even and odd parity including all boundary terms. *Class. Quantum Grav.*, **28**, 215017 [11 pp.].

Bañados, M., Teitelboim, C., and Zanelli, J. (1992). The black hole in three-dimensional space–time. *Phys. Rev. Lett.*, **69**, 1849–51.

Banerjee, R., Kiefer, C., and Majhi, B. R. (2010). Quantum gravitational corrections to the Hawking temperature from the Lemaître–Tolman–Bondi model. *Phys. Rev. D*, **82**, 044013 [7 pp.].

Banks, T. (1985). TCP, quantum gravity, the cosmological constant and all that. *Nucl. Phys. B*, **249**, 332–60.

Banks, T., Fischler, W., Shenker, S. H., and Susskind, L. (1997). M theory as a matrix model: a conjecture. *Phys. Rev. D*, **55**, 5112–28.

Barbero G., J. F. (1995). Real Ashtekar variables for Lorentzian space–times. *Phys. Rev. D*, **51**, 5507–10.

Barbero G., J. F. and Villaseñor, E. J. S. (2010). Quantization of midisuperspace models. *Living Rev. Relativ.*, **13**, article number 6.

Barbour, J. B. (1986). Leibnizian time, Machian dynamics, and quantum gravity. In *Quantum concepts in space and time* (ed. R. Penrose and C. J. Isham), pp. 236–46. Oxford University Press, Oxford.

Barbour, J. B. (1989). *Absolute or relative motion? Vol. 1: the discovery of dynamics.* Cambridge University Press, Cambridge.

Barbour, J. B. (1993). Time and complex numbers in canonical quantum gravity. *Phys. Rev. D*, **47**, 5422–9.

Barbour, J. B. (1994). The timelessness of quantum gravity: I. The evidence from the classical theory. *Class. Quantum Grav.*, **11**, 2853–73.

Barbour, J. B. (2000). *The end of time.* Oxford University Press, New York.

Barbour, J. B. (2009). The nature of time. arXiv:0903.3489v1 [gr-qc] [9 pp.].

Barbour, J. B. and Bertotti, B. (1982). Mach's principle and the structure of dynamical theories. *Proc. R. Soc. Lond. A*, **382**, 295–306.

Barbour, J. B. and Foster, B. (2008). Constraints and gauge transformations: Dirac's theorem is not always valid. arXiv:0808.1223v1 [gr-qc] [14 pp.].

Barbour, J. B., Foster, B., and Ó Murchadha, N. (2002). Relativity without relativity. *Class. Quantum Grav.*, **19**, 3217–48.

Barceló, C., Liberati, S., and Visser, M. (2011). Analogue gravity. *Living Rev. Relativ.*, **14**, article number 3.

Bardeen, J. M. (1980). Gauge invariant cosmological perturbations. *Phys. Rev. D*, **22**, 1882–905.

Bartnik, R. and Fodor, G. (1993). On the restricted validity of the thin sandwich conjecture. *Phys. Rev. D*, **48**, 3596–99.

Barvinsky, A. O. (1989). Perturbative quantum cosmology: the probability measure on superspace and semiclassical expansion. *Nucl. Phys. B*, **325**, 705–23.

Barvinsky, A. O. (1990). Effective action method in quantum field theory. In *Gauge theories of fundamental interactions* (ed. M. Pawlowski and R. Raczka), pp. 265–327. World Scientific, Singapore.

Barvinsky, A. O. (1993*a*). Unitarity approach to quantum cosmology. *Phys. Rep.*, **230**, 237–367.

Barvinsky, A. O. (1993*b*). Operator ordering in theories subject to constraints of the gravitational type. *Class. Quantum Grav.*, **10**, 1985–99.

Barvinsky, A. O. (1998). Solution of quantum Dirac constraints via path integral. *Nucl. Phys. B*, **520**, 533–60.

Barvinsky, A. O. (2001). Quantum cosmology at the turn of millennium. arXiv: gr-qc/0101046v1 [22 pp.].

Barvinsky, A. O. and Kamenshchik, A. Yu. (1990). One loop quantum cosmology: the normalizability of the Hartle–Hawking wave function and the probability of inflation. *Class. Quantum Grav.*, **7**, L181–6.

Barvinsky, A. O. and Kamenshchik, A. Yu. (1998). Effective equations of motion and initial conditions for inflation in quantum cosmology. *Nucl. Phys. B*, **532**, 339–60.

Barvinsky, A. O. and Kiefer, C. (1998). Wheeler–DeWitt equation and Feynman diagrams. *Nucl. Phys. B*, **526**, 509–39.

Barvinsky, A. O. and Krykhtin, V. (1993). Dirac and BFV quantization methods in the 1-loop approximation: closure of the quantum constraint algebra and the conserved inner product. *Class. Quantum Grav.*, **10**, 1957–84.

Barvinsky, A. O. and Kunstatter, G. (1996). Exact physical black hole states in generic 2-*D* dilaton gravity. *Phys. Lett. B*, **389**, 231–6.

Barvinsky, A. O. and Nesterov, D. V. (2001). Effective equations in quantum cosmology. *Nucl. Phys. B*, **608**, 333–74.

Barvinsky, A. O. and Vilkovisky, G. A. (1987). The effective action in quantum field theory: two-loop approximation. In *Quantum field theory and quantum statistics*, Vol. I (ed. I. Batalin, C. J. Isham, and G. A. Vilkovisky), pp. 245–75. Adam Hilger, Bristol.

Barvinsky, A. O., Kamenshchik, A. Yu., and Karmazin, I. P. (1993). The renormalization group for nonrenormalizable theories: Einstein gravity with a scalar field. *Phys. Rev. D*, **48**, 3677–94.

Barvinsky, A. O., Kamenshchik, A. Yu., and Mishakov, I. V. (1997). Quantum origin of the early inflationary universe. *Nucl. Phys. B*, **491**, 387–426.

Barvinsky, A. O., Kamenshchik, A. Yu., Kiefer, C., and Mishakov, I. V. (1999*a*). Decoherence in quantum cosmology at the onset of inflation. *Nucl. Phys. B*, **551**, 374–96.

Barvinsky, A. O., Kamenshchik, A. Yu., and Kiefer, C. (1999*b*). Effective action and decoherence by fermions in quantum cosmology. *Nucl. Phys. B*, **552**, 420–44.

Barvinsky, A. O., Kamenshchik, A. Yu., Rathke, A., and Kiefer, C. (2003*a*). Nonlocal braneworld action: an alternative to the Kaluza–Klein description. *Phys. Rev. D*, **67**, 023513 [19 pp.].

Barvinsky, A. O., Kamenshchik, A. Yu., Rathke, A., and Kiefer, C. (2003*b*). Radion-induced gravitational wave oscillations and their phenomenology. *Ann. Phys. (Berlin)*, 8th series, **12**, 343–70.

Barvinsky, A. O., Kamenshchik, A. Yu., Kiefer, C., Starobinsky, A. A., and Steinwachs, C. (2009). Asymptotic freedom in inflationary cosmology with a non-minimally coupled Higgs field. *J. Cosmol. Astropart. Phys.*, **12**, 003 [28 pp.].

Barvinsky, A. O., Kamenshchik, A. Yu., Kiefer, C., and Steinwachs, C. (2010). Tunneling cosmological state revisited: origin of inflation with a nonminimally coupled standard model Higgs inflaton. *Phys. Rev. D*, **81**, 043530 [9 pp.].

Batalin, I. A. and Fradkin, E. S. (1983). Operator quantization of relativistic dynamical systems subject to first class constraints. *Phys. Lett. B*, **128**, 303–8.

Batalin, I. A. and Vilkovisky, G. A. (1977). Relativistic S matrix of dynamical systems with boson and fermion constraints. *Phys. Lett. B*, **69**, 309–12.

Baumgarte, T. W. and Shapiro, S. L. (2003). Numerical relativity and compact binaries. *Phys. Rep.*, **376**, 41–131.

Becker, K., Becker, M., and Schwarz, J. H. (2007). *String theory and M-theory.* Cambridge University Press, Cambridge.

Beig, R. and Ó Murchadha, N. (1987). The Poincaré group as the symmetry group of canonical general relativity. *Ann. Phys. (NY)*, **174**, 463–98.

Bekenstein, J. D. (1973). Black holes and entropy. *Phys. Rev. D*, **7**, 2333–46.

Bekenstein, J. D. (1974). The quantum mass spectrum of the Kerr black hole. *Lett. Nuovo Cimento*, **11**, 467–70.

Bekenstein, J. D. (1999). Quantum black holes as atoms. In *Proc. 8th Marcel Grossmann Meet.* (ed. T. Piran and R. Ruffini), pp. 92–111. World Scientific, Singapore.

Bekenstein, J. D. (2001). The limits of information. *Stud. Hist. Philos. Mod. Phys.*, **32**, 511–24.

Bekenstein, J. D. (2004). Black holes and information theory. *Contemp. Phys.*, **45**, 31–43.

Bekenstein, J. D. and Mukhanov, V. F. (1995). Spectroscopy of the quantum black hole. *Phys. Lett. B*, **360**, 7–12.

Belinskii, V. A., Khalatnikov, I. M., and Lifshitz, E. M. (1982). A general solution of the Einstein equations with a time singularity. *Adv. Phys.*, **31**, 639–67.

Bell, J. S. (2004). *Speakable and unspeakable in quantum mechanics*, 2nd edn. Cambridge University Press, Cambridge.

Benedetti, D., Machado, P. F., and Saueressig, F. (2010). Taming perturbative divergences in asymptotically safe gravity. *Nucl. Phys. B*, **824**, 168–91.

Bento, M. C. and Bertolami, O. (1995). Scale factor duality: a quantum cosmological approach. *Class. Quantum Grav.*, **12**, 1919–26.

Berenstein, D., Maldacena, J., and Nastase, H. (2002). Strings in flat space and pp waves from $\mathcal{N} = 4$ super Yang Mills. *J. High Energy Phys.*, **04**, 013.

Berger, B. (1993). Quantum cosmology. In *Directions in general relativity, Vol. I* (ed. B. L. Hu, M. R. Ryan, and C. V. Vishveshwara), pp. 43–57. Cambridge University Press, Cambridge.

Berges, J., Tetradis, N., and Wetterich, C. (2002). Non-perturbative renormalization flow in quantum field theory and statistical physics. *Phys. Rep.*, **363**, 223–386.

Bergmann, P. G. (1961). Observables in general relativity. *Rev. Mod. Phys.*, **33**, 510–4.

Bergmann, P. G. (1966). Hamilton–Jacobi and Schrödinger theory in theories with first-class Hamiltonian constraints. *Phys. Rev.*, **144**, 1078–80.

Bergmann, P. G. (1989). The canonical formulation of general relativistic theories: the early years, 1930–1959. In *Einstein and the history of general relativity* (ed. D. Howard and J. Stachel), pp. 293–9. Birkhäuser, Boston.

Bergmann, P. G. (1992). Quantization of the gravitational field, 1930–1988. In *Studies in the history of general relativity* (ed. J. Eisenstaedt and A. J. Knox), pp. 364–6. Birkhäuser, Boston.

Bergmann, P. G. and Komar, A. (1960). Poisson brackets between locally defined observables in general relativity. *Phys. Rev. Lett.*, **4**, 432–3.

Bergmann, P. G. and Komar, A. (1972). The coordinate group symmetries of general relativity. *Int. J. Theor. Phys.*, **5**, 15–28.

Berkovits, N., Shen, A., and Zwiebach, B. (2000). Tachyon condensation in superstring field theory. *Nucl. Phys. B*, **587**, 147–78.

Bern, Z., Carrasco, J. J. M., Dixon, L. J., Johansson, H., and Roiban, R. (2009). Ultraviolet behavior of $\mathcal{N} = 8$ supergravity at four loops. *Phys. Rev. Lett.*, **103**, 081301 [4 pp.].

Berry, M. V. (1984). Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A*, **392**, 45–57.

Bertlmann, R. A. (1996). *Anomalies in quantum field theory.* Clarendon Press, Oxford.

Bertoni, C., Finelli, F., and Venturi, G. (1996). The Born–Oppenheimer approach to the matter–gravity system and unitarity. *Class. Quantum Grav.*, **13**, 2375–83.

Birrell, N. D. and Davies, P. C. W. (1982). *Quantum fields in curved space.* Cambridge University Press, Cambridge.

Bjerrum-Bohr, N. E. J., Donoghue, J. F., and Holstein, B. R. (2003*a*). Quantum gravitational corrections to the nonrelativistic scattering potential of two masses. *Phys. Rev. D*, **67**, 084033 [12 pp.].

Bjerrum-Bohr, N. E. J., Donoghue, J. F., and Holstein, B. R. (2003*b*). Quantum corrections to the Schwarzschild and Kerr metrics. *Phys. Rev. D*, **68**, 084005 [16 pp.].

Blais, D., Bringmann, T., Kiefer, C., and Polarski, D. (2003). Accurate results for primordial black holes from spectra with a distinguished scale. *Phys. Rev. D*, **67**, 024024 [11 pp.].

Blau, M. and Theisen, S. (2009). String theory as a theory of quantum gravity: a status report. *Gen. Relativ. Gravit.*, **41**, 743–55.

Błaut, A. and Kowalski-Glikman, J. (1998). The time evolution of quantum universe in the quantum potential picture. *Phys. Lett. A*, **245**, 197–202.

Blyth, W. F. and Isham, C. J. (1975). Quantization of a Friedmann universe filled with a scalar field. *Phys. Rev. D*, **11**, 768–78.

Böhm, M., Denner, A., and Joos, H. (2001). *Gauge theories of the strong and electroweak interaction.* Teubner, Stuttgart.

Bohr, N. (1949). Discussion with Einstein on epistemological problems in atomic physics. In *Albert Einstein: Philosopher-Scientist* (ed. P. A. Schilpp), pp. 200–41. Library of Living Philosophers, Vol. VII. Northwestern University, Evanston, Illinois.

Bohr, N. and Rosenfeld, L. (1933). Zur Frage der Messbarkeit der elektromagnetischen Feldgrössen. *Det Kgl. Danske Videnskabernes Selskab. Mathematisk - fysiske Meddelelser*, **XII**, 8, 3–65. English translation in Wheeler, J. A. and Zurek, W. H. (ed.), *Quantum theory and measurement*. Princeton University Press, Princeton (1983).

Bojowald, M. (2001*a*). Loop quantum cosmology IV: discrete time evolution. *Class. Quantum Grav.*, **18**, 1071–88.

Bojowald, M. (2001*b*). The semiclassical limit of loop quantum cosmology. *Class. Quantum Grav.*, **18**, L109–16.

Bojowald, M. (2004). Spherically symmetric quantum geometry: states and basic operators. *Class. Quantum Grav.*, **21**, 3733–53.

Bojowald, M. (2011). *Quantum cosmology*. Lecture Notes in Physics 835. Springer, Berlin.

Bojowald, M., Reyes, J. D., and Tibrewala, R. (2009). Nonmarginal Lemaitre–Tolman–Bondi-like models with inverse triad corrections from loop quantum gravity. *Phys. Rev. D*, **80**, 084002 [19 pp.].

Bojowald, M., Kiefer, C., and Moniz, P. V. (2010). Quantum cosmology for the 21st century: a debate. arXiv:1005.2471v1 [gr-qc] [20 pp.].

Bojowald, M., Calcagni, G., and Tsujikawa, S. (2011). Observational test of inflation in loop quantum cosmology. *J. Cosmol. Astropart. Phys.*, **11**, 046 [32 pp.].

Bonanno, A. and Reuter, R. (2006). Spacetime structure of an evaporating black hole in quantum gravity. *Phys. Rev. D*, **73**, 083005 [10 pp.].

Bonanno, A. and Reuter, R. (2011). Entropy production during asymptotically safe inflation. *Entropy*, **13**, 274–92.

Borde, A., Guth, A. H., and Vilenkin, A. (2003). Inflationary spacetimes are incomplete in past directions. *Phys. Rev. Lett.*, **90**, 151301 [4 pp.].

Börner, G. (2003). *The early universe: facts and fiction*, 4th edn. Springer, Berlin.

Boughn, S. and Rothman, T. (2006). Graviton emission and absorption by atomic hydrogen. *Class. Quantum Grav.*, **23**, 5839–52.

Bouhmadi-López, M. and Moniz, P. V. (2007). Quantization of parameters and the string landscape problem. *J. Cosmol. Astropart. Phys.*, **05**, 005.

Bouhmadi-López, M., Kiefer, C., Sandhöfer, B., and Moniz, P. V. (2009). Quantum fate of singularities in a dark-energy dominated universe. *Phys. Rev. D*, **79**, 124035 [15 pp.].

Boulanger, N., Damour, T., Gualtieri, L., and Henneaux, M. (2001). Inconsistency of interacting, multi-graviton theories. *Nucl. Phys. B*, **597**, 127–71.

Boulware, D. G. (1984). Quantization of higher derivative theories of gravity. In *Quantum theory of gravity* (ed. S. M. Christensen), pp. 267–94. Adam Hilger, Bristol.

Boulware, D. G. and Deser, S. (1967). Stress-tensor commutators and Schwinger terms. *J. Math. Phys.*, **8**, 1468–77.

Bousso, R. (2002). The holographic principle. *Rev. Mod. Phys.*, **74**, 825–74.

Bousso, R. and Polchinski, J. (2000). Quantization of four-form fluxes and dynamical neutralization of the cosmological constant. *J. High Energy Phys.*, **06**, 006.

Braun, J., Gies, H., and Scherer, D. D. (2011). Asymptotic safety: a simple example. *Phys. Rev. D*, **83**, 085012 [15 pp.].

Briggs, J. S. and Rost, J. M. (2001). On the derivation of the time-dependent equation of Schrödinger. *Found. Phys.*, **31**, 693–712.

Bronstein, M. (1936). Quantentheorie schwacher Gravitationsfelder. *Phys. Z. Sowjetunion*, **9**, 140–57. For an English translation, see *Gen. Relativ. Gravit.*, **44**, 267–83 (2012).

Brotz, T. and Kiefer, C. (1997). Semiclassical black hole states and entropy. *Phys. Rev. D*, **55**, 2186–91.

Brout, R. and Venturi, G. (1989). Time in semiclassical gravity. *Phys. Rev. D*, **39**, 2436–9.

Brown, J. D. (1988). *Lower dimensional gravity.* World Scientific, Singapore.

Brown, J. D. and Kuchař, K. V. (1995). Dust as a standard of space and time in canonical quantum gravity. *Phys. Rev. D*, **51**, 5600–29.

Brown, J. D. and York, J. W. (1989). Jacobi's action and the recovery of time in general relativity. *Phys. Rev. D*, **40**, 3312–18.

Brügmann, B. (1994). Loop representations. In *Canonical gravity: from classical to quantum* (ed. J. Ehlers and H. Friedrich), pp. 213–53. Lecture Notes in Physics 434. Springer, Berlin.

Buchbinder, I. L., Odintsov, S. D., and Shapiro, I. L. (1992). *Effective action in quantum gravity.* Institute of Physics Publishing, Bristol.

Burgess, C. P. (2004). Quantum gravity in everyday life: general relativity as an effective field theory. *Living Rev. Relativ.*, **7**, article number 5.

Butterfield, J. and Isham, C. J. (1999). On the emergence of time in quantum gravity. In *The arguments of time* (ed. J. Butterfield), pp. 111–68. Oxford University Press, Oxford.

Callan, C. G., Giddings, S. B., Harvey, S. A., and Strominger, A. (1992). Evanescent black holes. *Phys. Rev. D*, **45**, R1005–9.

Calmet, X. (2010). A review of quantum gravity at the Large Hadron Collider. *Mod. Phys. Lett. A*, **25**, 1553–79.

Candelas, P., Horowitz, G. T., Strominger, A., and Witten, E. (1985). Vacuum configurations for superstrings. *Nucl. Phys. B*, **258**, 46–74.

Cangemi, D., Jackiw, R., and Zwiebach, B. (1996). Physical states in matter-coupled dilaton gravity. *Ann. Phys. (NY)*, **245**, 408–44.

Capozziello, S. and Faraoni, V. (2011). *Beyond Einstein gravity.* Fundamental Theories of Physics 170. Springer, Dordrecht.

Carlip, S. (1998). *Quantum gravity in 2+1 dimensions.* Cambridge University Press, Cambridge.

Carlip, S. (2001). Quantum gravity: a progress report. *Rep. Prog. Phys.*, **64**, 885–942.

Carlip, S. (2008). Is quantum gravity necessary? *Class. Quantum Grav.*, **25**, 154010 [6 pp.].

Carlip, S. (2010). The small scale structure of spacetime. arXiv:1009.1136v1 [gr-qc] [14 pp.].

Carlip, S. and Teitelboim, C. (1995). The off-shell black hole. *Class. Quantum Grav.*, **12**, 1699–704.

Carr, B. J. (2003). Primordial black holes as a probe of cosmology and high energy physics. In *Quantum gravity: from theory to experimental search* (ed. D. Giulini, C. Kiefer, and C. Lämmerzahl), pp. 301–21. Lecture Notes in Physics 631. Springer, Berlin.

Carr, B. J. (ed.) (2007). *Universe or multiverse?* Cambridge University Press, Cambridge.

Carr, B. J., Kazunori, K., Sendouda, Y., and Yokoyama, J. (2010). New cosmological constraints on primordial black holes. *Phys. Rev. D*, **81**, 104019 [33 pp.].

Carroll, S. M., Friedman, D. Z., Ortiz, M. E., and Page, D. N. (1994). Physical states in canonically quantized supergravity. *Nucl. Phys. B*, **423**, 661–85.

Castellani, L. (1982). Symmetries in constrained Hamiltonian systems. *Ann. Phys. (NY)*, **143**, 357–71.

Cavaglià, M. and Moniz, P. V. (2001). Canonical and quantum FRW cosmological solutions in M-theory. *Class. Quantum Grav.*, **18**, 95–120.

Choquet-Bruhat, Y. and York, J. (1980). The Cauchy problem. In *General relativity and gravitation*, Vol. I (ed. A. Held), pp. 99–172. Plenum Press, New York.

Christodoulakis, T. and Korfiatis, E. (1991). Contact transformations and the quantisation of constraint systems. *Phys. Lett. B*, **256**, 484–90.

Christodoulakis, T. and Zanelli, J. (1987). Canonical approach to quantum gravity. *Class. Quantum Grav.*, **4**, 851–67.

Cline, D. B., Matthey, C., and Otwinowski, S. (2003). Evidence for a galactic origin of very short gamma ray bursts and primordial black hole sources. *Astropart. Phys.*, **18**, 531–8.

CMS Collaboration (2011). Search for microscopic black hole signatures at the Large Hadron Collider. *Phys. Lett. B*, **697**, 434–53.

Conradi, H. D. (1992). Initial state in quantum cosmology. *Phys. Rev. D*, **46**, 612–9.

Conradi, H. D. (1998). Tunneling of macroscopic universes. *Int. J. Mod. Phys. D*, **7**, 189–200.

Conradi, H. D. and Zeh, H. D. (1991). Quantum cosmology as an initial value problem. *Phys. Lett. A*, **154**, 321–6.

Coule, D. H. (2005). Quantum cosmological models. *Class. Quantum Grav.*, **22**, R125–66.

Crispino, L. C. B., Higuchi, A., and Matsas, G. E. A. (2008). The Unruh effect and its applications. *Rev. Mod. Phys.*, **80**, 787–838.

Csordás, A. and Graham, R. (1995). Exact quantum state for $N = 1$ supergravity. *Phys. Rev. D*, **52**, 6656–9.

Dąbrowski, M. P. and Kiefer, C. (1997). Boundary conditions in quantum string cosmology. *Phys. Lett. B*, **397**, 185–92.

Dąbrowski, M. P., Kiefer, C., and Sandhöfer, B. (2006). Quantum phantom cosmology. *Phys. Rev. D*, **74**, 044022 [12 pp.].

Damour, T. and Taylor, J. H. (1991). On the orbital period change of the binary pulsar PSR 1913+16. *Astrophys. J.*, **366**, 501–11.

Das, S. and Vagenas, E. C. (2008). Universality of quantum gravity corrections. *Phys. Rev. Lett.*, **101**, 221301 [4 pp.].

Das, S., Shankaranarayanan, S., and Sur, S. (2008). Black hole entropy from entanglement: a review. arXiv:0806.0402v1 [gr-qc] [28 pp.].

Dasgupta, A. and Loll, R. (2001). A proper-time cure for the conformal sickness in quantum gravity. *Nucl. Phys. B*, **606**, 357–79.

Davies, P. C. W. (1975). Scalar particle production in Schwarzschild and Rindler metrics. *J. Phys. A*, **8**, 609–16.

D'Eath, P. D. (1984). Canonical quantization of supergravity. *Phys. Rev. D*, **29**, 2199–219.

D'Eath, P. D. (1996). *Supersymmetric quantum cosmology.* Cambridge University Press, Cambridge.

de Calan, C. (1995). Construction of the Gross–Neveu model in dimension 3. In *Constructive physics* (ed. V. Rivasseau), pp. 149–59. Springer, Berlin.

Demers, J.-G. and Kiefer, C. (1996). Decoherence of black holes by Hawking radiation. *Phys. Rev. D*, **53**, 7050–61.

Deser, S. (1970). Self-interaction and gauge invariance. *Gen. Relativ. Gravit.* **1**, 9–18.

Deser, S. (1989). Quantum gravity: whence, whither. In *Trends in theoretical physics* (ed. P. J. Ellis and Y. C. Tang), pp. 175–91. Addison-Wesley, Reading, MA.

Deser, S. (2010). Gravity from self-interaction redux. *Gen. Relativ. Gravit.* **42**, 641–6.

Deser, S., van Nieuwenhuizen, P., and Boulware, D. (1975). Uniqueness and non-renormalisability of quantum gravitation. In *General relativity and gravitation* (ed. G. Shaviv and J. Rosen), pp. 1–18. Wiley, New York.

De Simone, A., Guth, A. H., Linde, A., Noorbala, M., Salem, M. P., and Vilenkin, A. (2010). Boltzmann brains and the scale-factor cutoff measure of the multiverse. *Phys. Rev. D*, **82**, 063520 [30 pp.].

de Wit, B. (2006). Supersymmetric black holes. *Fortschr. Phys.*, **54**, 183–98.

de Wit, B. and Nicolai, H. (1982). $N = 8$ supergravity. *Nucl. Phys. B*, **208**, 323–64.

de Wit, B., Hoppe, J., and Nicolai, H. (1988). On the quantum mechanics of supermembranes. *Nucl. Phys. B*, **305**, 545–81.

DeWitt, C. (ed.) (1957). *Proceedings of the conference on the role of gravitation in physics*, University of North Carolina, Chapel Hill, January 18–23, 1957. WADC Technical Report 57-216 (unpublished).

DeWitt, B. S. (1962). The quantization of geometry. In *Gravitation: an introduction to current research* (ed. L. Witten), pp. 266–381. Wiley, New York.

DeWitt, B. S. (1964). Gravity: a universal regulator? *Phys. Rev. Lett.*, **13**, 114–8.

DeWitt, B. S. (1965). *Dynamical theory of groups and fields.* Blackie & Son, London.

DeWitt, B. S. (1967*a*). Quantum theory of gravity. I. The canonical theory. *Phys. Rev.*, **160**, 1113–48.

DeWitt, B. S. (1967*b*). Quantum theory of gravity. II. The manifestly covariant theory. *Phys. Rev.*, **162**, 1195–239.

DeWitt, B. S. (1967*c*). Quantum theory of gravity. III. Applications of the covariant theory. *Phys. Rev.*, **162**, 1239–62.

DeWitt, B. S. (1970). Spacetime as a sheaf of geodesics in superspace. In *Relativity* (ed. M. Carmeli, S. I. Fickler, and L. Witten), pp. 359–74. Plenum Press, New York.

DeWitt, B. S. (1979). Quantum gravity: the new synthesis. In *General relativity: an Einstein centenary survey* (ed. S. W. Hawking and W. Israel), pp. 680–745.

Cambridge University Press, Cambridge.

DeWitt, B. S. (1999). The quantum and gravity: the Wheeler–DeWitt equation. In *The eighth Marcel Grossmann meeting* (ed. T. Piran and R. Ruffini), pp. 6–25. World Scientific, Singapore.

DeWitt, B. S. (2003). *The global approach to quantum field theory*, Vols. I and II. Clarendon Press, Oxford.

DeWitt-Morette, C. (2011). *The pursuit of quantum gravity.* Springer, Berlin.

Diakonov, D. (2011). Towards lattice-regularized quantum gravity. arXiv:1109.0091v1 [hep-th] [20 pp.].

Dimopoulos, S. and Landsberg, G. (2001). Black holes at the Large Hadron Collider. *Phys. Rev. Lett.*, **87**, 161602 [4 pp.].

Dimopoulos, S., Graham, P. W., Hogan, J. M., and Kasevich, M. A. (2007). Testing general relativity with atom interferometry. *Phys. Rev. Lett.*, **98**, 111102 [4 pp.].

Diósi, L. and Kiefer, C. (2000). Robustness and diffusion of pointer states. *Phys. Rev. Lett.*, **85**, 3552–5.

Diósi, L., Gisin, N., and Strunz, W. T. (2000). Quantum approach to coupling classical and quantum dynamics. *Phys. Rev. A*, **61**, 22108 [4 pp.].

Dirac, P. A. M. (1933). Homogeneous variables in classical dynamics. *Proc. Cambridge Philos. Soc.*, **29**, Part III, 389–400.

Dirac, P. A. M. (1958*a*). *The principles of quantum mechanics*, 4th edn. Clarendon Press, Oxford.

Dirac, P. A. M. (1958*b*). The theory of gravitation in Hamiltonian form. *Proc. R. Soc. A*, **246**, 333–43.

Dirac, P. A. M. (1959). Fixation of coordinates in the Hamiltonian theory of gravitation. *Phys. Rev.*, **114**, 924–30.

Dirac, P. A. M. (1964). *Lectures on quantum mechanics.* Belfer Graduate School of Science, Yeshiva University, New York.

Dirac, P. A. M. (1968). The quantization of the gravitational field. In *Contemporary physics: Trieste Symposium*, Vol. 1, pp. 539–43. IAEA, Vienna.

Dirac, P. A. M. (1981). Does renormalization make sense? In *Perturbative quantum chromodynamics* (ed. D. W. Duke and J. F. Owens), pp. 129–30. AIP Conference Proceedings, New York.

Dittrich, B. and Thiemann, T. (2009). Are the spectra of geometrical operators in loop quantum gravity really discrete? *J. Math. Phys.*, **50**, 012503 [11 pp.].

Domagala, M. and Lewandowski, J. (2004). Black-hole entropy from quantum geometry. *Class. Quantum Grav.*, **21**, 5233–43.

Donoghue, J. F. (1994). General relativity as an effective field theory. *Phys. Rev. D*, **50**, 3874–88.

Donoghue, J. F. and Torma, T. (1999). Infrared behavior of graviton–graviton scattering. *Phys. Rev. D*, **60**, 024003 [6 pp.].

Douglas, M. R. (2003). The statistics of string/M theory vacua. *J. High Energy Phys.*, **05**, 046.

Douglas, M. R. and Nekrasov, N. A. (2002). Noncommutative field theory. *Rev. Mod. Phys.*, **73**, 977–1029.

Dowker, F. (2006). Causal sets as discrete spacetime. *Contemp. Phys.*, **47**, 1–9.

Dreyer, O. (2003). Quasinormal modes, the area spectrum, and black hole entropy. *Phys. Rev. Lett.*, **90**, 081301 [4 pp.].

Duff, M. J. (1975). Covariant quantization. In *Quantum gravity: an Oxford symposium* (ed. C. J. Isham, R. Penrose, and D. W. Sciama), pp. 78–135. Clarendon Press, Oxford.

Duff, M. J., Okun, L. B., and Veneziano, G. (2002). Trialogue on the number of fundamental constants. *J. High Energy Phys.*, **03**, 023.

Dunne, G. V. (2005). Heisenberg–Euler effective Lagrangians: basics and extensions. In *From fields to strings: circumnavigating theoretical physics* (ed. M. Shifman, A. Vainshtein, and J. Wheater), pp. 445–522. World Scientific, Singapore.

Dyson, F. (2004). The world on a string. *New York Rev. Books*, **51**, Issue 8, 16–9.

Egan, C. A. and Lineweaver, C. H. (2010). A larger estimate of the entropy of the universe. *Astrophys. J.*, **710**, 1825–34.

Ehlers, J. (1973). The nature and structure of spacetime. In *The physicist's conception of nature* (ed. J. Mehra), pp. 71–91. Reidel, Dordrecht.

Ehlers, J. (1995). Machian ideas and general relativity. In *Mach's principle: from Newton's bucket to general relativity* (ed. J. B. Barbour and H. Pfister), pp. 458–73. Birkhäuser, Boston.

Einstein, A. (1916*a*). Hamiltonsches Prinzip und allgemeine Relativitätstheorie. *Sitzber. kgl.-preuß. Akad. Wiss. Berlin, Sitzung der phys.-math. Klasse*, **XLII**, 1111–6.

Einstein, A. (1916*b*). Näherungsweise Integration der Feldgleichungen der Gravitation. *Sitzber. kgl.-preuß. Akad. Wiss. Berlin, Sitzung der phys.-math. Klasse*, **XXXII**, 688–96.

Einstein, A. (1955). Preface in *Cinquant'anni di Relatività 1905–1955* (ed. M. Pantaleo). Coed. Giuntine e Sansoni, Firenze.

Einstein, A. and Fokker, A. D. (1914). Die Nordströmsche Gravitationstheorie vom Standpunkt des absoluten Differentialkalküls. *Ann. Phys. (Berlin)*, 4th series, **44**, 321–8.

Eppley, K. and Hannah, E. (1977). The necessity of quantizing the gravitational field. *Found. Phys.*, **7**, 51–68.

Esposito, G. (2011). An introduction to quantum gravity. arXiv:1108.3269 [58 pp.].

Everett, H. (1957). 'Relative state' formulation of quantum mechanics. *Rev. Mod. Phys.*, **29**, 454–62.

Faddeev, L. D. and Popov, V. N. (1967). Feynman diagrams for the Yang–Mills field. *Phys. Lett. B*, **25**, 29–30.

Faller, S. (2008). Effective field theory of gravity: leading quantum gravitational corrections to Newton's and Coulomb's law. *Phys. Rev. D*, **77**, 124039 [14 pp.].

Farrugia, Ch. J. and Hajicek, P. (1979). The third law of black hole mechanics: a counterexample. *Commun. Math. Phys.*, **68**, 291–9.

Fels, M. E. and Torre, C. G. (2002). The principle of symmetric criticality in general relativity. *Class. Quantum Grav.*, **19**, 641–76.

Feng, J. L., March-Russell, J., Sethi, S., and Wilczek, F. (2001). Saltatory relaxation of the cosmological constant. *Nucl. Phys. B*, **602**, 307–28.

Fertig, C. and Gibble, K. (2000). Measurement and cancellation of the cold collision frequency in an ^{87}Rb fountain clock. *Phys. Rev. Lett.*, **85**, 1622–5.

Feynman, R. P. (1963). Quantum theory of gravitation. *Acta Phys. Polonica*, **XXIV**, 697–722.

Feynman, R. P. and Hibbs, A. R. (1965). *Quantum mechanics and path integrals.* McGraw Hill, Boston.

Feynman, R. P., Morinigo, F. B., and Wagner, W. G. (1995). *Feynman lectures on gravitation.* Addison-Wesley, Reading, MA.

Fierz, M. and Pauli, W. (1939). On relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc. R. Soc. A*, **173**, 211–32.

Fischer, A. E. (1970). The theory of superspace. In *Relativity* (ed. M. Carmeli, S. I. Fickler, and L. Witten), pp. 303–57. Plenum Press, New York.

Fischer, A. E. (1986). Resolving the singularities in the space of Riemannian geometries. *J. Math. Phys.*, **27**, 718–38.

Fleischhack, C. (2009). Representations of the Weyl algebra in quantum geometry. *Commun. Math. Phys.*, **285**, 67–140.

Ford, L. H. (1982). Gravitational radiation by quantum systems. *Ann. Phys. (NY)*, **144**, 238–48.

Ford, L. H. (2005). Spacetime in semiclassical gravity. arXiv:gr-qc/0504096v2 [18 pp.].

Fradkin, E. S. and Vasiliev, M. A. (1977). Hamiltonian formalism, quantization and S-matrix for supergravity. *Phys. Lett. B*, **72**, 70–4.

Franzen, A., Gutti, S., and Kiefer, C. (2010). Quantum gravitational collapse in the Lemaitre–Tolman–Bondi mode with a positive cosmological constant. *Class. Quantum Grav.*, **27**, 015011 [12 pp.].

Fray, S., Alvarez Diez, C., Hänsch, T. W., and Weitz, M. (2004). Atomic interferometer with amplitude gratings of light and its applications to atom based tests of the equivalence principle. *Phys. Rev. Lett.*, **93**, 240404 [4 pp.].

Fré, P., Gorini, V., Magli, G., and Moschella, U. (eds.) (1999). *Classical and quantum black holes.* Institute of Physics Publishing, Bristol.

Fredenhagen, K. and Haag, R. (1990). On the derivation of Hawking radiation associated with the formation of a black hole. *Commun. Math. Phys.*, **127**, 273–84.

Freidel, L., Livine, E. R., and Rovelli, C. (2003). Spectra of length and area in 2+1 Lorentzian loop quantum gravity. *Class. Quantum Grav.*, **20**, 1463–78.

Friedman, J. L. and Sorkin, R. D. (1980). Spin 1/2 from gravity. *Phys. Rev. Lett.*, **44**, 1100–3.

Frieman, J., Brandenberger, R., Kiefer, C., Müller, V., Mukhanov, V., Sato, K., *et al.* (1997). Are we making progress in relating cosmology and fundamental theories? In *The evolution of the Universe* (ed. G. Börner and S. Gottlöber), pp. 141–56. Wiley, Chichester.

Frittelli, S., Lehner, L., and Rovelli, C. (1996). The complete spectrum of the area from recoupling theory in loop quantum gravity. *Class. Quantum Grav.*, **13**, 2921–32.

Frolov, V. P. and Novikov, I. D. (1998). *Black hole physics.* Kluwer, Dordrecht.

Frolov, V. P. and Vilkovisky, G. A. (1981). Spherical symmetric collapse in quantum gravity. *Phys. Lett. B*, **106**, 307–13.

Fuji, Y. and Maeda, K. (2003). *The scalar–tensor theory of gravitation.* Cambridge University Press, Cambridge.

Fulling, S. A. (1973). Nonuniqueness of canonical field quantization in Riemannian space–time. *Phys. Rev. D*, **7**, 2850–62.

Fulling, S. A. (1989). *Aspects of quantum field theory in curved space–time.* Cambridge University Press, Cambridge.

Gambini, R. and Pullin, J. (2011). *A first course in loop quantum gravity.* Oxford University Press, Oxford.

Gamow, G., Ivanenko, D., and Landau, L. (2002). World constants and limiting transition. *Phys. At. Nuclei*, **65**, 1373–5.

Garay, L. J. (1995). Quantum gravity and minimal length. *Int. J. Mod. Phys. A*, **10**, 145–66.

Gasperini, M. and Veneziano, G. (2003). The pre-big bang scenario in string cosmology. *Phys. Rep.*, **373**, 1–212.

Gasser, J. and Leutwyler, H. (1984). Chiral perturbation theory to one loop. *Ann. Phys. (NY)*, **158**, 142–210.

Gerhardt, C. (2009). Quantum cosmological Friedman models with an initial singularity. *Class. Quantum Grav.*, **26**, 015001 [29 pp.].

Gerlach, U. H. (1969). Derivation of the ten Einstein field equations from the semiclassical approximation to quantum geometrodynamics. *Phys. Rev.*, **177**, 1929–41.

Geroch, R. and Hartle, J. B. (1986). Computability and physical theories. *Found. Phys.*, **16**, 533–50.

Gibbons, G. W. and Hartle, J. B. (1990). Real tunneling geometries and the large-scale topology of the universe. *Phys. Rev. D*, **42**, 2458–68.

Gibbons, G. W. and Hawking, S. W. (1977). Cosmological event horizons, thermodynamics, and particle creation. *Phys. Rev. D*, **15**, 2738–51.

Gibbons, G. W., Hawking, S. W., and Perry, M. J. (1978). Path integrals and the indefiniteness of the gravitational action. *Nucl. Phys. B*, **138**, 141–50.

Giddings, S. B. (2011). Is string theory a theory of quantum gravity? arXiv: 1105.-6359v1 [hep-th] [32 pp.].

Giddings, S. B. and Thomas, S. (2002). High energy colliders as black hole factories: the end of short distance physics. *Phys. Rev. D*, **65**, 056010 [12 pp.].

Giles, R. (1981). Reconstruction of gauge potentials from Wilson loops. *Phys. Rev. D*, **24**, 2160–8.

Giulini, D. (1995*a*). On the configuration space topology in general relativity. *Helv. Phys. Acta*, **68**, 86–111.

Giulini, D. (1995*b*). What is the geometry of superspace? *Phys. Rev. D*, **51**, 5630–5.

Giulini, D. (1999). The generalized thin-sandwich problem and its local solvability. *J. Math. Phys.*, **40**, 2470–82.

Giulini, D. (2003). That strange procedure called quantisation. In *Quantum gravity: from theory to experimental search* (ed. D. Giulini, C. Kiefer, and C. Lämmerzahl), pp. 17–40. Lecture Notes in Physics 631. Springer, Berlin.

Giulini, D. (2007). Remarks on the notions of general covariance and background independence. In *Approaches to fundamental physics: an assessment of current theoretical ideas* (ed. E. Seiler and I.-O. Stamatescu), pp. 105–20. Lecture Notes in Physics 721. Springer, Berlin.

Giulini, D. (2009). The superspace of geometrodynamics. *Gen. Relativ. Gravit.*, **41**, 785–815.

Giulini, D. and Großardt, A. (2011). Gravitationally induced inhibitions of dispersion according to the Schrödinger–Newton equation. *Class. Quantum Grav.*, 195026 [17 pp.].

Giulini, D. and Kiefer, C. (1994). Wheeler–DeWitt metric and the attractivity of gravity. *Phys. Lett. A*, **193**, 21–4.

Giulini, D. and Kiefer, C. (1995). Consistency of semiclassical gravity. *Class. Quantum Grav.*, **12**, 403–11.

Giulini, D. and Kiefer, C. (2007). The canonical approach to quantum gravity: general ideas and geometrodynamics. In *Approaches to fundamental physics: an assessment of current theoretical ideas* (ed. E. Seiler and I.-O. Stamatescu), pp. 131–50. Lecture Notes in Physics 721. Springer, Berlin.

Giulini, D. and Marolf, D. (1999). A uniqueness theorem for constraint quantization. *Class. Quantum Grav.*, **16**, 2489–505.

Goldhaber, A. S. and Nieto, M. N. (2010). Photon and graviton mass limits. *Rev. Mod. Phys.*, **82**, 939–79.

Gorelik, G. E. (2005). Matvei Bronstein and quantum gravity: 70th anniversary of the unsolved problem. *Phys. Uspekhi*, **48**(10), 1039–53.

Goroff, M. H. and Sagnotti, A. (1985). Quantum gravity at two loops. *Phys. Lett. B*, **160**, 81–6.

Gousheh, S. S. and Sepangi, H. R. (2000). Wave packets and initial conditions in quantum cosmology. *Phys. Lett. A*, **272**, 304–12.

Green, M. B., Schwarz, J. H., and Witten, E. (1987). *Superstring theory*, 2 vols. Cambridge University Press, Cambridge.

Greene, B. (1997). String theory on Calabi–Yau manifolds. arXiv:hep-th/9702155v1 [150 pp.].

Grib, A. A., Mamayev, S. G., and Mostepanenko, V. M. (1994). *Vacuum effects in strong fields*. Friedmann Laboratory Publishing, St Petersburg.

Grishchuk, L. P. and Sidorov, Y. V. (1990). Squeezed quantum states of relic gravitons and primordial density fluctuations. *Phys. Rev. D*, **42**, 3413–21.

Gronwald, F. and Hehl, F. W. (1996). On the gauge aspects of gravity. arXiv:gr-qc/9602013v1 [50 pp.].

Gross, D. and Periwal, V. (1988). String perturbation theory diverges. *Phys. Rev. Lett.*, **60**, 2105–8.

Grumiller, D., Kummer, W., and Vassilevich, D. V. (2002). Dilaton gravity in two dimensions. *Phys. Rep.*, **369**, 327–430.

Guica, M., Hartman, T., Song, W., and Strominger, A. (2009). The Kerr/CFT correspondence. *Phys. Rev. D*, **80**, 124008 [9 pp.].

Gupta, S. N. (1952*a*). Quantization of Einstein's gravitational field: linear approximation. *Proc. R. Soc. A*, **65**, 161–9.

Gupta, S. N. (1952*b*). Quantization of Einstein's gravitational field: general treatment. *Proc. R. Soc. A*, **65**, 608–19.

Haag, R., Lopuszanski, J. T., and Sohnius, M. (1975). All possible generators of supersymmetries of the S matrix. *Nucl. Phys. B*, **88**, 257–74.

Hájíček, P. (2001). Unitary dynamics of spherical null gravitating shells. *Nucl. Phys. B*, **603**, 555–77.

Hájíček, P. (2003). Quantum theory of gravitational collapse (lecture notes on quantum conchology). In *Quantum gravity: from theory to experimental search* (ed. D. Giulini, C. Kiefer, and C. Lämmerzahl), pp. 255–99. Lecture Notes in Physics 631. Springer, Berlin.

Hájíček, P. and Kiefer, C. (2001*a*). Embedding variables in the canonical theory of gravitating shells. *Nucl. Phys. B*, **603**, 531–54.

Hájíček, P. and Kiefer, C. (2001*b*). Singularity avoidance by collapsing shells in quantum gravity. *Int. J. Mod. Phys. D*, **10**, 775–9.

Hájíček, P. and Kijowski, J. (2000). Covariant gauge fixing and Kuchař decomposition. *Phys. Rev. D*, **61**, 024037 [13 pp.].

Hall, M. J. W. and Reginatto, M. (2005). Interacting classical and quantum ensembles. *Phys. Rev. A*, **72**, 062109 [13 pp.].

Halliwell, J. J. (1988). Derivation of the Wheeler–DeWitt equation from a path integral for minisuperspace models. *Phys. Rev. D*, **38**, 2468–81.

Halliwell, J. J. (1989). Decoherence in quantum cosmology. *Phys. Rev. D*, **39**, 2912–23.

Halliwell, J. J. (1991). Introductory lectures on quantum cosmology. In *Quantum cosmology and baby universes* (ed. S. Coleman, J. B. Hartle, T. Piran, and S. Weinberg), pp. 159–243. World Scientific, Singapore.

Halliwell, J. J. (1998). Effective theories of coupled classical and quantum variables from decoherent histories: a new approach to the back reaction problem. *Phys. Rev. D*, **57**, 2337–48.

Halliwell, J. J. and Hartle, J. B. (1991). Wave functions constructed from an invariant sum over histories satisfy constraints. *Phys. Rev. D*, **43**, 1170–94.

Halliwell, J. J. and Hawking, S. W. (1985). Origin of structure in the Universe. *Phys. Rev. D*, **31**, 1777–91.

Halliwell, J. J. and Louko, J. (1990). Steepest descent contours in the path integral approach to quantum cosmology. 3. A general method with applications to anisotropic minisuperspace models. *Phys. Rev. D*, **42**, 3997–4031.

Hamber, H. W. (2009). *Quantum gravitation: the Feynman path integral approach.* Springer, Berlin.

Hanson, A. J., Regge, T., and Teitelboim, C. (1976). *Constrained Hamiltonian systems.* Accademia Nazionale dei Lincei, Roma.

Harrison, E. (2000). *Cosmology*, 2nd edn. Cambridge University Press, Cambridge.

Hartle, J. B. (1987). Prediction in quantum cosmology. In *Gravitation in astrophysics* (ed. B. Carter and J. B. Hartle), pp. 329–60. Plenum Press, New York.

Hartle, J. B. (1997). Quantum cosmology: problems for the 21st century. In *Physics in the 21st century* (ed. K. Kikkawa, H. Kunitomo, and H. Ohtsubo), pp. 179–99. World Scientific, Singapore.

Hartle, J. B. and Hawking, S. W. (1976). Path-integral derivation of black-hole radiance. *Phys. Rev. D*, **13**, 2188–203.

Hartle, J. B. and Hawking, S. W. (1983). Wave function of the Universe. *Phys. Rev. D*, **28**, 2960–75.

Hartle, J. B. and Schleich, K. (1987). The conformal rotation in linearised gravity. In *Quantum field theory and quantum statistics*, Vol. II (ed. I. Batalin, C. J. Isham, and G. A. Vilkovisky), pp. 67–87. Adam Hilger, Bristol.

Hawking, S. W. (1975). Particle creation by black holes. *Commun. Math. Phys.*, **43**, 199–220. Erratum *ibid.* **46**, 206 (1976).

Hawking, S. W. (1976). Breakdown of predictability in gravitational collapse. *Phys. Rev. D*, **14**, 2460–73.

Hawking, S. W. (1979). The path-integral approach to quantum gravity. In *General relativity: an Einstein centenary survey* (ed. S. W. Hawking and W. Israel), pp. 746–89. Cambridge University Press, Cambridge.

Hawking, S. W. (1982). The boundary conditions of the universe. *Pontif. Acad. Sci. Scr. Varia*, **48**, 563–74.

Hawking, S. W. (1984). The quantum state of the universe. *Nucl. Phys. B*, **239**, 257–76.

Hawking, S. W. (2005). Information loss in black holes. *Phys. Rev. D*, **72**, 084013 [4 pp.].

Hawking, S. W. and Ellis, G. F. R. (1973). *The large scale structure of space–time.* Cambridge University Press, Cambridge.

Hawking, S. W. and Penrose, R. (1996). *The nature of space and time.* Princeton University Press, Princeton, NJ.

Hayward, G. (1993). Gravitational action for spacetimes with nonsmooth boundaries. *Phys. Rev. D*, **47**, 3275–80.

Hehl, F. W. (1985). On the kinematics of the torsion of space–time. *Found. Phys.*, **15**, 451–71.

Hehl, F. W., von der Heyde, P., Kerlick, G. D., and Nester, J. M. (1976). General relativity with spin and torsion. *Rev. Mod. Phys.*, **48**, 393–416.

Hehl, F. W., Lemke, J., and Mielke, E.W. (1991). Two lectures on fermions and gravity. In *Geometry and theoretical physics* (ed. J. Debrus and A. C. Hirshfeld), pp. 56–140. Springer, Berlin.

Hehl, F. W., Kiefer, C., and Metzler, R. J. K. (eds.) (1998). *Black holes: theory and observation.* Lecture Notes in Physics 514. Springer, Berlin.

Heisenberg, W. (1985). *Collected works*, Abt. C, Vol. III (ed. H.-P. Dürr and H. Rechenberg). Piper, Munich.

Heitler, W. (1984). *The quantum theory of radiation.* Dover Publications, New York.

Helfer, A. D. (1996). The stress-energy operator. *Class. Quantum Grav.*, **13**, L129–34.

Henneaux, M. and Teitelboim, C. (1992). *Quantization of gauge systems.* Princeton University Press, Princeton, NJ.

Heusler, M. (1996). *Black hole uniqueness theorems.* Cambridge University Press, Cambridge.

Heusler, M., Kiefer, C., and Straumann, N. (1990). Self-energy of a thin charged shell in general relativity. *Phys. Rev. D*, **42**, 4254–6.

Higgs, P. W. (1958). Integration of secondary constraints in quantized general relativity. *Phys. Rev. Lett.*, **1**, 373–4.

Hod, S. (1998). Bohr's correspondence principle and the area spectrum of quantum black holes. *Phys. Rev. Lett.*, **81**, 4293–6.

Hogan, C. J. (2000). Why the universe is just so. *Rev. Mod. Phys.*, **72**, 1149–61.
Hogan, C. J. (2002). Holographic discreteness of inflationary perturbations. *Phys. Rev. D*, **66**, 023521 [9 pp.].
Hojman, S. A., Kuchař, K. V., and Teitelboim, C. (1976). Geometrodynamics regained. *Ann. Phys. (NY)*, **96**, 88–135.
Hořava, P. (2009). Quantum gravity at a Lifshitz point. *Phys. Rev. D*, **79**, 084008 [15 pp.].
Hornberger, K., Uttenthaler, S., Brezger, B., Hackermueller, L., Arndt, M., and Zeilinger, A. (2003). Collisional decoherence observed in matter wave interferometry. *Phys. Rev. Lett.*, **90**, 160401 [4 pp.].
Horowitz, G. T. (1990). String theory as a quantum theory of gravity. In *General relativity and gravitation* (ed. N. Ashby, D. F. Bartlett, and W. Wyss), pp. 419–39. Cambridge University Press, Cambridge.
Horowitz, G. T. (1998). Quantum states of black holes. In *Black holes and relativistic stars* (ed. R. M. Wald), pp. 241–66. University of Chicago Press, Chicago.
Horowitz, G. T. (2005). Spacetime in string theory. *New J. Phys.*, **7**, 201 [13 pp.].
Horowitz, G. T. and Polchinski, J. (1997). Correspondence principle for black holes and strings. *Phys. Rev. D*, **55**, 6189–97.
Hossenfelder, S. (2010). Experimental search for quantum gravity. arXiv:1010. 3420v1 [gr-qc] [24 pp.].
Hsu, S. D. H. and Reeb, D. (2009). Black holes, information, and decoherence. *Phys. Rev. D*, **79**, 124037 [7 pp.].
Hu, B. L. and Verdaguer, E. (2004). Stochastic gravity: theory and applications. *Living Rev. Relativ.*, **7**, article number 3.
Huang, K. (1992). *Quarks, leptons & gauge fields*, 2nd edn. World Scientific, Singapore.
Husain, V. and Winkler, O. (2004). Singularity resolution in quantum gravity. *Phys. Rev. D*, **69**, 084016 [7 pp.].
Husain, V. and Winkler, O. (2005). Quantum resolution of black hole singularities. *Class. Quantum Grav.*, **22**, L127–33.
Immirzi, G. (1997). Quantum gravity and Regge calculus. *Nucl. Phys. B (Proc. Suppl.)*, **57**, 65–72.
Infeld, L. (ed.) (1964). *Relativistic theories of gravitation.* Pergamon Press, Oxford.
Isham, C. J. (1976). Some quantum field theory aspects of the superspace quantization of general relativity. *Proc. R. Soc. Lond. A*, **351**, 209–32.
Isham, C. J. (1981). Topological θ-sectors in canonically quantized gravity. *Phys. Lett. B*, **106**, 188–92.
Isham, C. J. (1984). Topological and global aspects of quantum theory. In *Relativity, groups and topology II* (ed. B. S. DeWitt and R. Stora), pp. 1059–290. North-Holland, Amsterdam.
Isham, C. J. (1987). Quantum gravity. In *General relativity and gravitation* (ed. M. A. H. MacCallum), pp. 99–129. Cambridge University Press, Cambridge.
Isham, C. J. (1989). Quantum topology and quantization on the lattice of topologies. *Class. Quantum Grav.*, **6**, 1509–34.
Isham, C. J. (1993). Canonical quantum gravity and the problem of time. In *Integrable*

systems, quantum groups, and quantum field theory (ed. L. A. Ibort and M. A. Rodríguez), pp. 157–287. Kluwer, Dordrecht.

Isham, C. J. (1994). Prima facie questions in quantum gravity. In *Canonical gravity: from classical to quantum* (ed. J. Ehlers and H. Friedrich), pp. 1–21. Lecture Notes in Physics 434. Springer, Berlin.

Israel, W. (1986). Third law of black-hole dynamics: a formulation and proof. *Phys. Rev. Lett.*, **57**, 397–99.

Iwasaki, Y. (1971). Quantum theory of gravitation vs. classical theory. *Prog. Theor. Phys.*, **46**, 1587–609.

Jackiw, R. (1995). *Diverse topics in theoretical and mathematical physics.* World Scientific, Singapore.

Jackson, J. D. (2008). Examples of the zeroth theorem of the history of science. *Am. J. Phys.*, **76**, 704–19.

Jacobson, T. (1995). Thermodynamics of spacetime: the Einstein equation of state. *Phys. Rev. Lett.*, **75**, 1260–3.

Jacobson, T. (2003). Introduction to quantum fields in curved spacetime and the Hawking effect. arXiv:gr-qc/0308048v3 [59 pp.].

Joos, E., Zeh, H. D., Kiefer, C., Giulini, D., Kupsch, J., and Stamatescu, I.-O. (2003). *Decoherence and the appearance of a classical world in quantum theory*, 2nd edn. Springer, Berlin.

Kaku, M. (1999). *Introduction to superstrings and M-theory*, 2nd edn. Springer, New York.

Kamenshchik, A. Yu. (1990). Normalizability of the wave function of the universe, particle physics and supersymmetry. *Phys. Lett. B*, **316**, 45–50.

Kamenshchik, A. Y., Kiefer, C., and Sandhöfer, B. (2007). Quantum cosmology with a big-brake singularity. *Phys. Rev. D*, **76**, 064032 [13 pp.].

Kanti, P. (2004). Black holes in theories with large extra dimensions: a review. *Int. J. Mod. Phys. A*, **19**, 4899–951.

Kastrup, H. A. (1996). The quantum levels of isolated spherically symmetric gravitational systems. *Phys. Lett. B*, **385**, 75–80.

Kastrup, H. A. (2008). On the advancements of conformal transformations and their associated symmetries in geometry and theoretical physics. *Ann. Phys. (Berlin)*, 8th series, **17**, 631–90.

Kastrup, H. A. and Thiemann, T. (1994). Spherically symmetric gravity as a completely integrable system. *Nucl. Phys. B*, **425**, 665–86.

Kazakov, D. I. (1988). On a generalization of renormalization group equations to quantum field theories of an arbitrary type. *Theor. Math. Phys.*, **75**, 440–2.

Kiefer, C. (1987). Continuous measurement of minisuperspace variables by higher multipoles. *Class. Quantum Grav.*, **4**, 1369–82.

Kiefer, C. (1988). Wave packets in minisuperspace. *Phys. Rev. D*, **38**, 1761–72.

Kiefer, C. (1989). Non-minimally coupled scalar fields and the initial value problem in quantum gravity. *Phys. Lett. B*, **225**, 227–32.

Kiefer, C. (1990). Wave packets in quantum cosmology and the cosmological constant. *Nucl. Phys. B*, **341**, 273–93.

Kiefer, C. (1991). On the meaning of path integrals in quantum cosmology. *Ann. Phys. (NY)*, **207**, 53–70.

Kiefer, C. (1992). Decoherence in quantum electrodynamics and quantum gravity. *Phys. Rev. D*, **46**, 1658–70.

Kiefer, C. (1994). The semiclassical approximation to quantum gravity. In *Canonical gravity: from classical to quantum* (ed. J. Ehlers and H. Friedrich), pp. 170–212. Lecture Notes in Physics 434. Springer, Berlin.

Kiefer, C. (1996). Quantum gravitational effects in de Sitter space. In *New frontiers in gravitation* (ed. G. A. Sardanashvily), pp. 203–14. Hadronic Press, Palm Harbor. See arXiv:gr-qc/9501001v1 for a related version.

Kiefer, C. (1998). Towards a full quantum theory of black holes. In *Black holes: theory and observation* (ed. F. W. Hehl, C. Kiefer, and R. J. K. Metzler), pp. 416–50. Lecture Notes in Physics 514. Springer, Berlin.

Kiefer, C. (1999). Thermodynamics of black holes and Hawking radiation. In *Classical and quantum black holes* (ed. P. Fré, V. Gorini, G. Magli, and U. Moschella), pp. 17–74. Institute of Physics Publishing, Bristol.

Kiefer, C. (2001*a*). Path integrals in quantum cosmology. In *Fluctuating paths and fields* (ed. W. Janke, A. Pelster, H.-J. Schmidt, and M. Bachmann), pp. 729–40. World Scientific, Singapore.

Kiefer, C. (2001*b*). Hawking radiation from decoherence. *Class. Quantum Grav.*, **18**, L151–4.

Kiefer, C. (2003*a*). Quantum aspects of black holes. In *The galactic black hole* (ed. H. Falcke and F. W. Hehl), pp. 207–25. Institute of Physics Publishing, Bristol.

Kiefer, C. (2003*b*). On the interpretation of quantum theory—from Copenhagen to the present day. In *Time, quantum and information* (ed. L. Castell and O. Ischebeck), pp. 291–9. Springer, Berlin.

Kiefer, C. (2004*a*). Is there an information-loss problem for black holes? In *Decoherence and entropy in complex systems* (ed. H.-T. Elze). Lecture Notes in Physics 633. Springer, Berlin.

Kiefer, C. (2004*b*). Hawking temperature from quasi-normal modes. *Class. Quantum Grav.*, **21**, L123–7.

Kiefer, C. (2006). Quantum gravity: general introduction and recent developments. *Ann. Phys. (Berlin)*, 8th series, **15**, 129–48.

Kiefer, C. (2009). Can the arrow of time be understood from quantum cosmology? arXiv:0910.5836v1 [gr-qc] [14 pp.].

Kiefer, C. (2010). Can singularities be avoided in quantum cosmology? *Ann. Phys. (Berlin)*, 8th series, **19**, 211–8.

Kiefer, C. and Kolland, G. (2008). Gibbs' paradox and black-hole entropy. *Gen. Relativ. Gravit.*, **40**, 1327–39.

Kiefer, C. and Krämer, M. (2012). Quantum gravitational contributions to the CMB anisotropy spectrum. *Phys. Rev. Lett.*, **108**, 021301 [4 pp.].

Kiefer, C. and Louko, J. (1999). Hamiltonian evolution and quantization for extremal black holes. *Ann. Phys. (Berlin)*, 8th series, **8**, 67–81.

Kiefer, C. and Singh, T. P. (1991). Quantum gravitational correction terms to the functional Schrödinger equation. *Phys. Rev. D*, **44**, 1067–76.

Kiefer, C. and Wipf, A. (1994). Functional Schrödinger equation for fermions in external gauge fields. *Ann. Phys. (NY)*, **236**, 241–85.

Kiefer, C. and Weber, C. (2005). On the interaction of mesoscopic quantum systems with gravity. *Ann. Phys. (Berlin)*, 8th series, **14**, 253–78.

Kiefer, C. and Zeh, H. D. (1995). Arrow of time in a recollapsing quantum universe. *Phys. Rev. D*, **51**, 4145–53.

Kiefer, C., Polarski, D., and Starobinsky, A. A. (1998). Quantum-to-classical transition for fluctuations in the early universe. *Int. J. Mod. Phys. D*, **7**, 455–62.

Kiefer, C., Polarski, D., and Starobinsky, A. A. (2000). Entropy of gravitons produced in the early universe. *Phys. Rev. D*, **62**, 043518 [6 pp.].

Kiefer, C., Lück, T., and Moniz, P. (2005). Semiclassical approximation to supersymmetric quantum gravity. *Phys. Rev. D*, **72**, 045006 [19 pp.].

Kiefer, C., Müller-Hill, J., and Vaz, C. (2006). Classical and quantum Lemaître–Tolman–Bondi model for the nonmarginal case. *Phys. Rev. D*, **73**, 044025 [13 pp.]. Typographical errors are corrected in arXiv:gr-qc/0512047v2.

Kiefer, C., Lohmar, I., Polarski, D., and Starobinsky, A. A. (2007*a*). Pointer states for primordial fluctuations in inflationary cosmology. *Class. Quantum Grav.*, **24**, 1699–1718.

Kiefer, C., Müller-Hill, J., Singh, T. P., and Vaz, C. (2007*b*). Hawking radiation from the quantum Lemaître–Tolman–Bondi model. *Phys. Rev. D*, **75**, 124010 [8 pp.].

Kiefer, C., Marto, J., and Moniz, P. V. (2009). Indefinite oscillators and black-hole evaporation. *Ann. Phys. (Berlin)*, 8th series, **18**, 722–35.

Kiefer, C., Queisser, F., and Starobinsky, A. A. (2011). Cosmological constant from decoherence. *Class. Quantum Grav.*, **28**, 125022 [26 pp.].

Klauder, J. R. (2001). Quantization of constrained systems. In *Methods of quantization* (ed. H. Latal and W. Schweiger), pp. 143–83. Lecture Notes in Physics 572. Springer, Berlin.

Klauder, J. R. (2010). An affinity for affine quantum gravity. arXiv:1003.2617v1 [gr-qc] [14 pp.].

Kleinert, H. (2009). *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*, 5th edn. World Scientific, Singapore.

Kleinschmidt, A., Koehn, M., and Nicolai, H. (2009). Supersymmetric quantum cosmological billiards. *Phys. Rev. D*, **80**, 061701(R) [5 pp.].

Kodama, H. (1990). Holomorphic wavefunction of the universe. *Phys. Rev. D*, **42**, 2548–65.

Kokkotas, K. D. and Schmidt, B. G. (1999). Quasi-normal modes of stars and black holes. *Living Rev. Relativ.*, **2**, article number 2.

Komar, A. (1979). Constraints, hermiticity, and correspondence. *Phys. Rev. D*, **19**, 2908–12.

Kontoleon, N. and Wiltshire, D. L. (1996). Operator ordering and consistency of the wave function of the Universe. *Phys. Rev. D*, **59**, 063513 [8 pp.].

Kostelecký, V. A. and Russell, N. (2011). Data tables for Lorentz and CPT violation. *Rev. Mod. Phys.*, **83**, 11–31.

Krasnov, K. (2010). Effective metric Lagrangians from an underlying theory with two propagating degrees of freedom. *Phys. Rev. D*, **81**, 084026 [40 pp.].

Kraus, P. and Wilczek, F. (1995). Self-interaction correction to black hole radiance. *Nucl. Phys. B*, **433**, 403–20.

Kretschmann, E. (1917). Über den physikalischen Sinn der Relativitätspostulate A. Einsteins und seine ursprüngliche Relativitätstheorie. *Ann. Phys. (Berlin)*, 4th series, **53**, 575–614.

Kuchař, K. (1970). Ground state functional of the linearized gravitational field. *J. Math. Phys.*, **11**, 3322–34.

Kuchař, K. (1971). Canonical quantization of cylindrical gravitational waves. *Phys. Rev. D*, **4**, 955–86.

Kuchař, K. (1973). Canonical quantization of gravity. In *Relativity, astrophysics and cosmology* (ed. W. Israel), pp. 237–88. Reidel, Dordrecht.

Kuchař, K. (1981). Canonical methods of quantization. In *Quantum Gravity 2. A second Oxford symposium* (ed. C. J. Isham, R. Penrose, and D. W. Sciama), pp. 329–76. Oxford University Press, Oxford.

Kuchař, K. V. (1992). Time and interpretations of quantum gravity. In *Proc. 4th Canadian Conf. General Relativity Relativistic Astrophysics* (ed. G. Kunstatter, D. Vincent, and J. Williams), pp. 211–314. World Scientific, Singapore.

Kuchař, K. V. (1993). Canonical quantum gravity. In *General relativity and gravitation 1992* (ed. R. J. Gleiser, C. N. Kozameh, and O. M. Moreschi), pp. 119–50. Institute of Physics Publishing, Bristol.

Kuchař, K. V. (1994). Geometrodynamics of Schwarzschild black holes. *Phys. Rev. D*, **50**, 3961–81.

Kuchař, K. V. and Ryan, M. P. (1989). Is minisuperspace quantization valid?: Taub in mixmaster. *Phys. Rev. D*, **40**, 3982–96.

Kuchař, K. V. and Torre, C. G. (1991). Strings as poor relatives of general relativity. In *Conceptual problems of quantum gravity* (ed. A. Ashtekar and J. Stachel), pp. 326–48. Birkhäuser, Boston.

Kuo, C.-I. and Ford, L. H. (1993). Semiclassical gravity theory and quantum fluctuations. *Phys. Rev. D*, **47**, 4510–9.

Laflamme, R. (1987). Euclidean vacuum: justification from quantum cosmology. *Phys. Lett. B*, **198**, 156–60.

Laflamme, R. and Louko, J. (1991). Reduced density matrices and decoherence in quantum cosmology. *Phys. Rev. D*, **43**, 3317–31.

Lämmerzahl, C. (1995). A Hamilton operator for quantum optics in gravitational fields. *Phys. Lett. A*, **203**, 12–17.

Lämmerzahl, C. (1996). On the equivalence principle in quantum theory. *Gen. Relativ. Gravit.*, **28**, 1043–70.

Lämmerzahl, C. (1997). Constraints on space–time torsion from Hughes–Drever experiments. *Phys. Lett. A*, **228**, 223–31.

Lämmerzahl, C. (1998). Quantum tests of the foundations of general relativity. *Class. Quantum Grav.*, **15**, 13–27.

Lämmerzahl, C. (2003). The Einstein equivalence principle and the search for new physics. In *Quantum gravity: from theory to experimental search* (ed. D. Giulini, C. Kiefer, and C. Lämmerzahl), pp. 367–400. Lecture Notes in Physics 631. Springer, Berlin.

Lanczos, C. (1986). *The variational principles of mechanics*, 4th edn. Dover Publications, New York.

Landau, L. and Peierls, R. (1931). Erweiterung des Unbestimmtheitsprinzips für die relativistische Quantentheorie. *Z. Phys.*, **69**, 56–69.

Lapchinsky, V. G. and Rubakov, V. A. (1979). Canonical quantization of gravity and quantum field theory in curved space–time. *Acta Phys. Polonica*, **10**, 1041–8.

Lauscher, O. and Reuter, M. (2001). Ultraviolet fixed point and generalized flow equation of quantum gravity. *Phys. Rev. D*, **65**, 025013 [44 pp.].

Lauscher, O. and Reuter, M. (2002). Flow equation of quantum Einstein gravity in a higher-derivative truncation. *Phys. Rev. D*, **66**, 025026 [50 pp.].

Lee, H. C. (ed.) (1984). *An introduction to Kaluza–Klein theories.* World Scientific, Singapore.

Leinaas, J. M. (2002). Unruh effect in storage rings. In *Quantum aspects of beam physics* (ed. P. Chen), pp. 336–52. World Scientific, Singapore.

Lemaître, G. (1931). The beginning of the world from the point of view of quantum theory. *Nature*, **127**, 706. See J.-P. Luminet, arXiv:1105.6271v1 for an editorial note on this paper.

Lemaître, G. (1933). L'univers en expansion. *Ann. Soc. Sci. de Bruxelles*, **A53**, 51–85. For an English translation, see *Gen. Relativ. Gravit.*, **29**, 641–80 (1997).

Lemaître, G. (1958). The primaeval atom hypothesis and the problem of the clusters of galaxies. In *La structure et l'évolution de l'univers* (ed. R. Stoops), pp. 1–31. Coudenberg, Brussels.

Leutwyler, H. (1964). Gravitational field: equivalence of Feynman quantization and canonical quantization. *Phys. Rev.*, **134**, B1155–82.

Leutwyler, H. (1986). Anomalies. *Helv. Phys. Acta*, **59**, 201–19.

Lewandowski, J., Newman, E. T., and Rovelli, C. (1993). Variations of the parallel propagator and holonomy operator and the Gauss law constraint. *J. Math. Phys.*, **34**, 4646–54.

Lewandowski, J., Okolów, A., Sahlmann, H., and Thiemann, T. (2006). Uniqueness of diffeomorphism invariant states on holonomy-flux algebras. *Commun. Math. Phys.*, **267**, 703–33.

Liddle, A. R. and Lyth, D. H. (2000). *Cosmological inflation and large-scale structure.* Cambridge University Press, Cambridge.

Lidsey, J. E. (1995). Scale factor duality and hidden supersymmetry in scalar–tensor cosmology. *Phys. Rev. D*, **52**, 5407–11.

Lifshits, E. M. (1946). About gravitational stability of the expanding world [in Russian]. *Zh. Eksp. Teor. Phys.*, **16**, 587–602.

Linde, A. D. (1990). *Particle physics and inflationary cosmology.* Harwood, New York.

Linde, A. D. and Vanchurin, V. (2010). How many universes are in the multiverse? *Phys. Rev. D*, **81**, 083525 [11 pp.].

Loos, H. G. (1969). Canonical gauge- and Lorentz-invariant quantization of the Yang–Mills field. *Phys. Rev.*, **188**, 2342–53.

Louis-Martinez, D. and Kunstatter, G. (1994). Birkhoff's theorem in two-dimensional dilaton gravity. *Phys. Rev. D*, **49**, 5227–30.

Louis-Martinez, D., Gegenberg, J., and Kunstatter, G. (1994). Exact Dirac quantization of all 2D dilaton gravity theories. *Phys. Lett. B*, **321**, 193–8.

Louko, J. (2006). Group averaging, positive definiteness and superselection sectors. *J. Phys. Conf. Ser.* **33**, 142–9.

Louko, J. and Matschull, H.-J. (2001). The 2+1 Kepler problem and its quantization. *Class. Quantum Grav.*, **18**, 2731–84.

Louko, J. and Whiting, B. F. (1995). Hamiltonian thermodynamics of the Schwarzschild black hole. *Phys. Rev. D*, **51**, 5583–99.

Louko, J. and Winters-Hilt, S. N. (1996). Hamiltonian thermodynamics of the Reissner–Nordström–anti-de Sitter black hole. *Phys. Rev. D*, **54**, 2647–63.

Louko, J., Whiting, B., and Friedman, J. (1998). Hamiltonian spacetime dynamics with a spherical null-dust shell. *Phys. Rev. D*, **57**, 2279–98.

Lüscher, M., Narayanan, R., Weisz, P., and Wolff, U. (1992). The Schrödinger functional: a renormalizable probe for non-abelian gauge theories. *Nucl. Phys. B*, **384**, 168–228.

Lüst, D. and Theisen, S. (1989). *Lectures on string theory.* Lecture Notes in Physics 346. Springer, Berlin.

Maartens, R. and Koyama, K. (2010). Brane-world gravity. *Living Rev. Relativ.*, **13**, article number 5.

MacCallum, M. A. H. (1979). Anisotropic and inhomogeneous relativistic cosmologies. In *General relativity: an Einstein centenary survey* (ed. S. W. Hawking and W. Israel), pp. 533–80. Cambridge University Press, Cambridge.

MacGibbon, J. H. (1991). Quark- and gluon-jet emission from primordial black holes. II. The emission over the black-hole lifetime. *Phys. Rev. D*, **44**, 376–92.

MacGibbon, J. H. and Webber, B. R. (1990). Quark- and gluon-jet emission from primordial black holes: the instantaneous spectra. *Phys. Rev. D*, **41**, 3052–79.

MacGibbon, J. H., Carr, B. J., and Page, D. N. (2008). Do evaporating black holes form photospheres? *Phys. Rev. D*, **78**, 064043 [21 pp.].

Maggiore, M. (2008). *Gravitational waves, Vol. 1: theory and experiments.* Oxford University Press, Oxford.

Major, S. A. (1999). A spin network primer. *Am. J. Phys.*, **67**, 972–80.

Maldacena, J. (2011). The gauge/gravity duality. arXiv:1106.6073v1 [hep-th] [23 pp.].

Mandelstam, S. (1962). Quantization of the gravitational field. *Ann. Phys. (NY)*, **19**, 25–66 (1962).

Manrique, E., Rechenberger, S., and Saueressig, F. (2011*a*). Asymptotically safe Lorentzian gravity. *Phys. Rev. Lett.*, **106**, 251302 [4 pp.].

Manrique, E., Reuter, M., and Saueressig, F. (2011*b*). Bimetric renormalization group flows in quantum Einstein gravity. *Ann. Phys. (NY)*, **326**, 463–85.

Marsden, J. E. and Hughes, T. J. R. (1983). *Mathematical foundations of elasticity.* Prentice-Hall, Englewood Cliffs, New Jersey.

Martin, S. P. (1997). A supersymmetry primer. arXiv: hep-th/9709356v5 [128 pp.].

Mashhoon, B. (1995). On the coupling of intrinsic spin with the rotation of the earth. *Phys. Lett A*, **198**, 9–13.

Mathur, S. D. (2011). What the information paradox is *not.* arXiv:1108.0302v2 [hep-th] [40 pp.].

Matschull, H.-J. (1995). Three-dimensional canonical quantum gravity. *Class. Quantum Grav.*, **12**, 2621–703.

Matschull, H.-J. (2001). The phase space structure of multi-particle models in 2+1 gravity. *Class. Quantum Grav.*, **17**, 3497–560.

McAvity, D. M. and Osborn, H. (1993). Quantum field theories on manifolds with curved boundaries: scalar fields. *Nucl. Phys. B*, **394**, 728–88.

McGuigan, M. (1988). Third quantization and the Wheeler–DeWitt equation. *Phys. Rev. D*, **38**, 3031–51.

Meissner, K. A. (2004). Black-hole entropy in loop quantum gravity. *Class. Quantum Grav.*, **21**, 5245–51.

Mersini-Houghton, L. (2005). Can we predict Λ for the non-SUSY sector of the landscape? *Class. Quantum Grav.*, **22**, 3481–90.

Misner, C. W. (1957). Feynman quantization of general relativity. *Rev. Mod. Phys.*, **29**, 497–509.

Misner, C. W. (1972). Minisuperspace. In *Magic without magic: John Archibald Wheeler* (ed. J. R. Klauder), pp. 441–73. Freeman, San Francisco.

Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973). *Gravitation*. Freeman, San Francisco.

Modesto, L. (2004). Disappearance of the black hole singularity in loop quantum gravity. *Phys. Rev. D*, **70**, 124009 [5 pp.].

Modesto, L. (2009). Fractal spacetime from the area spectrum. *Class. Quantum Grav.*, **26**, 242002 [9 pp.].

Mohaupt, T. (2001). Black hole entropy, special geometry and strings. *Fortschr. Phys.*, **49**, 3–161.

Mohaupt, T. (2003). Introduction to string theory. In *Quantum gravity: from theory to experimental search* (ed. D. Giulini, C. Kiefer, and C. Lämmerzahl), pp. 173–251. Lecture Notes in Physics 631. Springer, Berlin.

Møller, C. (1962). The energy–momentum complex in general relativity and related problems. In *Les théories relativistes de la gravitation* (ed. A. Lichnerowicz and M. A. Tonnelat), pp. 15–29. Editions du Centre National de la Recherche Scientifique, Paris.

Moniz, P. V. (2003). FRW minisuperspace with local $N = 4$ supersymmetry and self-interacting scalar field. *Ann. Phys. (Berlin)*, 8th series, **12**, 174–98.

Moniz, P. V. (2010). *Quantum cosmology: the supersymmetric perspective*, Vols. 1 and 2. Lecture Notes in Physics 804. Springer, Berlin.

Motl, L. (2003). An analytical computation of asymptotic Schwarzschild quasinormal frequencies. *Adv. Theor. Math. Phys.*, **6**, 1135–62.

Mott, N. F. (1931). On the theory of excitation by collision with heavy particles. *Proc. Cambridge Philos. Soc.*, **27**, 553–60.

Mukhanov, V. F. (1986). Are black holes quantized? *JETP Lett.*, **44**, 63–6.

Myers, R. C. and Perry, M. J. (1986). Black holes in higher dimensional space-times. *Ann. Phys. (NY)*, **172**, 304–47.

Nakamura, K. and Particle Data Group (2010). Review of particle physics. *J. Phys. G: Nucl. Part. Phys.*, **37**, 075021 [1422 pp.].

Neitzke, A. (2003). Greybody factors at large imaginary frequencies. arXiv:hep-th/0304080v1 [19 pp.].

Nesvizhevsky, V. V., Börner, H. G., Petukhov, A. K., Abele, H., Baeßler, S., Rueß, F. J., Stöferle, T., Westphal, A., Gagarski, A. M., Petrov, G. A., and Strelkov, A. V. (2002). Quantum states of neutrons in the Earth's gravitational field. *Nature*, **415**, 297–9.

Nesvizhevsky, V. V., Voronin, A. Yu., Cubitt, R., and Protasov, K. V. (2010). Neutron whispering gallery. *Nature Phys.*, **6**, 114–7.

Neugebauer, G. (1998). Black hole thermodynamics. In *Black holes: theory and observation* (ed. F. W. Hehl, C. Kiefer, and R. J. K. Metzler), pp. 319–38. Lecture Notes in Physics 514. Springer, Berlin.

Nicolai, H. (1999). On M-theory. *J. Astrophys. Astron.*, **20**, 149–64.

Nicolai, H. and Matschull, H.-J. (1993). Aspects of canonical gravity and supergravity. *J. Geom. Phys.*, **11**, 15–62.

Nicolai, H., Peeters, K., and Zamaklar, M. (2005). Loop quantum gravity: an outside view. *Class. Quantum Grav.*, **22**, R193–R247.

Niemeyer, J. and Jedamzik, K. (1999). Dynamics of primordial black hole formation. *Phys. Rev. D*, **59**, 124013 [8 pp.].

Noether, E. (1918). Invariante Variationsprobleme. *Nachr. Ges. Wiss. Göttingen*, **1918**, 235–57.

Núñez, D., Quevedo, H., and Sudarsky, D. (1998). Black hole hair: a review. In *Black holes: theory and observation* (ed. F. W. Hehl, C. Kiefer, and R. J. K. Metzler), pp. 187–98. Lecture Notes in Physics 514. Springer, Berlin.

Okun, L. B. (1992). The fundamental constants of physics. *Sov. Phys. Usp.*, **34**, 818–26.

Oriti, D. (ed.) (2009). *Approaches to quantum gravity.* Cambridge University Press, Cambridge.

Osterwalder, K. and Schrader, R. (1975). Axioms for euclidean Green's functions II. *Commun. Math. Phys.*, **42**, 281–305.

Padmanabhan, T. (1985). Physical significance of Planck length. *Ann. Phys. (NY)*, **165**, 38–58.

Padmanabhan, T. (2010). Thermodynamical aspects of gravity: new insights. *Rep. Prog. Phys.*, **73**, 046901 [44 pp.].

Page, D. N. (1976). Particle emission rates from a black hole: massless particles from an uncharged, nonrotating hole. *Phys. Rev. D*, **13**, 198–206.

Page, D. N. (1991). Minisuperspaces with conformally and minimally coupled scalar fields. *J. Math. Phys.*, **32**, 3427–38.

Page, D. N. (1994). Black hole information. In *Proceedings of the 5th Canadian conference on general relativity and relativistic astrophysics* (ed. R. Mann and R. McLenaghan), pp. 1–41. World Scientific, Singapore.

Page, D. N. (2005). Hawking radiation and black hole thermodynamics. *New J. Phys.*, **7**, 203 [32 pp.].

Page, D. N. and Geilker, C. D. (1981). Indirect evidence for quantum gravity. *Phys. Rev. Lett.*, **47**, 979–82.

Page, D. N. and Hawking, S. W. (1976). Gamma rays from primordial black holes. *Astrophys. J.*, **206**, 1–7.

Palais, R. S. (1979). The principle of symmetric criticality. *Commun. Math. Phys.*, **69**, 19–30.

Parentani, R. (2000). The background field approximation in (quantum) cosmology. *Class. Quantum Grav.*, **17**, 1527–47.

Parikh, M. K. and Wilczek, F. (2000). Hawking radiation as tunneling. *Phys. Rev. Lett.*, **85**, 5042–5.

Parker, L. E. and Toms, D. J. (2009). *Quantum field theory in curved spacetime.* Cambridge University Press, Cambridge.

Paternoga, R. and Graham, R. (1998). The Chern–Simons state for the non-diagonal Bianchi IX model. *Phys. Rev. D*, **58**, 083501 [22 pp.].

Pauli, W. (1955). Schlußwort durch den Präsidenten der Konferenz. *Helv. Phys. Acta Suppl.*, **4**, 261–7.

Pauli, W. (1985). *Scientific correspondence with Bohr, Einstein, Heisenberg a.o., Vol. II.* Springer, Berlin.

Pauli, W. (1993). *Scientific correspondence with Bohr, Einstein, Heisenberg a.o., Vol .III*, p. 748. Springer, Berlin.

Paz, J. P. and Sinha, S. (1992). Decoherence and back reaction in quantum cosmology: multidimensional minisuperspace examples. *Phys. Rev. D*, **45**, 2823–42.

Peet, A. W. (1998). The Bekenstein formula and string theory (N-brane theory). *Class. Quantum Grav.*, **15**, 3291–338.

Penrose, R. (1971). Angular momentum: an approach to combinatorial space–time. In *Quantum theory and beyond* (ed. T. Bastin), pp. 151–80. Cambridge University Press, Cambridge.

Penrose, R. (1975). Twistor theory, its aims and achievements. In *Quantum gravity* (ed. C. J. Isham, R. Penrose, and D. W. Sciama), pp. 268–407. Clarendon Press, Oxford.

Penrose, R. (1981). Time-asymmetry and quantum gravity. In *Quantum gravity*, Vol. 2 (ed. C. J. Isham, R. Penrose, and D. W. Sciama), pp. 242–72. Clarendon Press, Oxford.

Penrose, R. (1996). On gravity's role in quantum state reduction. *Gen. Relativ. Gravit.*, **28**, 581–600.

Penrose, R. (2005). *The road to reality: a complete guide to the laws of the Universe.* Knopf, New York.

Percacci, R. (2009). Asymptotic safety. In *Approaches to quantum gravity* (ed. D. Oriti). pp. 111–28. Cambridge University Press, Cambridge.

Percival, I. and Strunz, W. (1997). Detection of space-time fluctuations by a model matter interferometer. *Proc. R. Soc. Lond. A*, **453**, 431–46.

Peres, A. (1962). On Cauchy's problem in general relativity—II. *Nuovo Cimento*, **XXVI**, 53–62.

Peres, A. (1995). *Quantum theory: concepts and methods.* Kluwer, Dordrecht.

Perez, A. (2006). The spin-foam representation of LQG. arXiv:gr-qc/0601095v1 [19 pp.].

Peters, A., Chung, K. Y., and Chu, S. (2001). High-precision gravity measurements using atom interferometry. *Metrologia*, **38**, 25–61.

Pilati, M. (1977). The canonical formulation of supergravity. *Nucl. Phys. B*, **132**, 138–54.

Pilati, M. (1982). Strong coupling quantum gravity. I. Solution in a particular gauge. *Phys. Rev. D*, **26**, 2645–63.

Pilati, M. (1983). Strong coupling quantum gravity. II. Solution without gauge fixing. *Phys. Rev. D*, **28**, 729–44.

Pinto-Neto, N. and Santini, E. S. (2002). The consistency of causal quantum geometrodynamics and quantum field theory. *Gen. Relativ. Gravit.*, **34**, 505–32.

Plaga, R. (2006). A fundamental threat to quantum cryptography: gravitational attacks. *Eur. Phys. J. D*, **38**, 409–13.

Planck, M. (1899). Über irreversible Strahlungsvorgänge. *Sitzber. kgl.-preuß. Akad. Wiss. Berlin, Sitzungen der phys.-math. Klasse*, pp. 440–80.

Plebański, J. and Krasiński, A. (2006). *An introduction to general relativity and cosmology.* Cambridge University Press, Cambridge.

Poincaré, H. (1970). *La valeur de la science.* Flammarion, Paris.

Poincaré, H. (2009). *La science et l'hypothèse.* Flammarion, Paris.

Polarski, D. and Starobinsky, A. A. (1996). Semiclassicality and decoherence of cosmological fluctuations. *Class. Quantum Grav.*, **13**, 377–92.

Polchinski, J. (1998*a*). *String theory. Vol. I. An introduction to the bosonic string.* Cambridge University Press, Cambridge.

Polchinski, J. (1998*b*). *String theory. Vol. II. Superstring theory and beyond.* Cambridge University Press, Cambridge.

Pons, J. M. (2005). On Dirac's incomplete analysis of gauge transformations. *Stud. Hist. Philos. Mod. Phys.*, **36**, 491–518.

Pons, J. M., Salisbury, D. C., and Sundermeyer, K. A. (2009). Revisiting observables in generally covariant theories in the light of gauge fixing methods. *Phys. Rev. D*, **80**, 084015 [23 pp.].

Randall, L. and Sundrum, R. (1999). A large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, **83**, 3370–3.

Rees, M. (1995). *Perspectives in astrophysical cosmology.* Cambridge University Press, Cambridge.

Regge, T. (1961). General relativity without coordinates. *Nuovo Cimento*, **XIX**, 558–71.

Regge, T. and Teitelboim, C. (1974). Role of surface integrals in the Hamiltonian formulation of general relativity. *Ann. Phys. (NY)*, **88**, 286–318.

Regge, T. and Williams, R. M. (2000). Discrete structures in gravity. *J. Math. Phys.*, **41**, 3964–84.

Reuter, M. (1998). Nonperturbative evolution equation for quantum gravity. *Phys. Rev. D*, **57**, 971–85.

Reuter, M. and Saueressig, F. (2002*a*). Renormalization group flow of quantum gravity in the Einstein–Hilbert truncation. *Phys. Rev. D*, **65**, 065016 [26 pp.].

Reuter, M. and Saueressig, F. (2002*b*). A class of nonlocal truncations in quantum Einstein gravity and its renormalization group behavior. *Phys. Rev. D*, **66**, 125001

[34 pp.].

Reuter, M. and Weyer, H. (2004). Quantum gravity at astrophysical distances?. *J. Cosmol. Astropart. Phys.*, **12**, 001.

Rickles, D. and French, S. (2006). Quantum gravity meets structuralism: interweaving relations in the foundations of physics. In *The structural foundations of quantum gravity* (ed. D. Rickles, S. French, and J. Saatsi), pp. 1–39. Oxford University Press, Oxford.

Rosenfeld, L. (1930*a*). Über die Gravitationswirkungen des Lichtes. *Z. Phys.*, **65**, 589–99.

Rosenfeld, L. (1930*b*). Zur Quantelung der Wellenfelder. *Ann. Phys. (Berlin)*, **397**, 113–52. For an English translation with commentary, see D. Salisbury, Max Planck Institute for the History of Science, Preprint 381 (2009).

Rosenfeld, L. (1963). On quantization of fields. *Nucl. Phys.*, **40**, 353–6.

Rovelli, C. (1991*a*). Ashtekar formulation of general relativity and loop-space non-perturbative quantum gravity: a report. *Class. Quantum Grav.*, **8**, 1613–75.

Rovelli, C. (1991*b*). Time in quantum gravity: an hypothesis. *Phys. Rev. D*, **43**, 442–56.

Rovelli, C. (2004). *Quantum gravity.* Cambridge University Press, Cambridge.

Rovelli, C. (2007). Comment on "Are the spectra of geometrical operators in loop quantum gravity really discrete?" by B. Dittrich and T. Thiemann. arXiv:0708.2481 [6 pp.].

Rovelli, C. (2011). A new look at loop quantum gravity. *Class. Quantum Grav.*, **28**, 114005 [24 pp.].

Rovelli, C. and Gaul, M. (2000). Loop quantum gravity and the meaning of diffeomorphism invariance. In *Towards quantum gravity* (ed. J. Kowalski-Glikman), pp. 277–324. Lecture Notes in Physics 541. Springer, Berlin.

Rovelli, C. and Smolin, L. (1990). Loop space representation of quantum general relativity. *Nucl. Phys. B*, **331**, 80–152.

Rovelli, C. and Smolin, L. (1995). Discreteness of area and volume in quantum gravity. *Nucl. Phys. B*, **442**, 593–622. Erratum *ibid.*, **456**, 753–4.

Rubakov, V. A. (2001). Large and infinite extra dimensions. *Phys. Usp.*, **44**, 871–93.

Ryan, M. P. (1972). *Hamiltonian cosmology.* Lecture Notes in Physics 13. Springer, Berlin.

Ryan, M. P. and Waller, S. M. (1997). On the Hamiltonian formulation of class B Bianchi cosmological models. arXiv:gr-qc/9709012v1 [21 pp.].

Sahlmann, H. (2011). Loop quantum gravity—a short review. arXiv:1001.4188v3 [gr-qc] [23 pp.].

Sakellariadou, M. (2010). Phenomenology of loop quantum cosmology. *J. Phys: Conf. Ser.*, **222**, 012027 [16 pp.].

Sannan, S. (1986). Gravity as the limit of the type-II superstring theory. *Phys. Rev. D*, **34**, 1749–58.

Scherk, J. and Schwarz, J. H. (1974). Dual models for non-hadrons. *Nucl. Phys. B*, **81**, 118–44.

Schlosshauer, M. (2006). Experimental motivation and empirical consistency in minimal no-collapse quantum mechanics. *Ann. Phys. (NY)*, **321**, 112–49.

Schlosshauer, M. (2007). *Decoherence and the quantum-to-classical transition.* Springer, Berlin.

Schmidt, H.-J. (1990). The metric in the superspace of Riemannian metrics and its relation to gravity. In *Differential geometry and its applications* (ed. J. Janyška and D. Kupka), pp. 405–11. World Scientific, Singapore.

Schrödinger, E. (1939). The proper vibrations of the expanding universe. *Physica*, **6**, 899–912.

Schuller, F. and Wohlfarth, M. N. R. (2006). Geometry of manifolds with area metric. *Nucl. Phys. B*, **747**, 398–422.

Schwinger, J. (1963). Quantized gravitational field. *Phys. Rev.*, **130**, 1253–8.

Sen, A. (1982). Gravity as a spin system. *Phys. Lett. B*, **119**, 89–91.

Sen, A. (1999). Universality of the tachyon potential. *J. High Energy Phys.*, **12**, 027.

Sexl, R. U. and Urbantke, H. K. (2001). *Relativity, groups, particles.* Springer, Vienna.

Shaposhnikov, M. and Wetterich, C. (2010). Asymptotic safety of gravity and the Higgs boson mass. *Phys. Lett. B*, **683**, 196–200.

Shi, Y. (2000). Early Gedanken experiments of quantum mechanics revisited. *Ann. Phys. (Berlin)*, 8th series, **9**, 637–48.

Singh, T. P. (2005). Quantum mechanics without spacetime: a case for noncommutative geometry. arXiv:gr-qc/0510042v1 [34 pp.].

Singh, T. P. and Padmanabhan, T. (1989). Notes on semiclassical gravity. *Ann. Phys. (NY)*, **196**, 296–344.

Smith, G. J. and Bergmann, P. G. (1979). Measurability analysis of the magnetic-type components of the linearized gravitational radiation field. *Gen. Relativ. Gravit.*, **11**, 133–47.

Solomon, J. (1938). Gravitation et quanta. *J. de Physique et de Radium*, **9**, 479–85.

Sorkin, R. (1997). Forks in the road, on the way to quantum gravity. *Int. J. Theor. Phys.*, **36**, 2759–81.

Sorkin, R. (2005). Causal sets: discrete gravity. In *Lectures on quantum gravity: proceedings of the Valdivia summer school* (ed. A. Gomberoff and D. Marolf), pp. 305–27. Plenum, New York.

Sotiriou, T. P. (2011). Hořava–Lifshitz gravity: a status report. *J. Phys.: Conf. Ser.*, **283**, 012034 [17 pp.].

Sotiriou, T. P., Visser, M., and Weinfurtner, S. (2011). Spectral dimension as a probe of the ultraviolet continuum regime of causal dynamical triangulations. arXiv: 1105.5646v1 [gr-qc] [5 pp.].

Stachel, J. (1999). The early history of quantum gravity (1916–1940). In *Black holes, gravitational radiation and the universe* (ed. B. R. Iyer and B. Bhawal), pp. 525–34. Kluwer, Dordrecht.

Starobinsky, A. A. (1979). Spectrum of relict gravitational radiation and the early state of the universe. *JETP Lett.*, **30**, 682–5.

Starobinsky, A. A. (1992). Spectrum of adiabatic perturbations in the universe when there are singularities in the inflation potential. *JETP Lett.*, **55**, 489–94.

Steinacker, H. (2010). Emergent geometry and gravity from matrix models: an introduction. *Class. Quantum Grav.*, **27**, 133001 [46 pp.].

Stelle, K. S. (1977). Renormalization of higher-derivative quantum gravity. *Phys. Rev. D*, **16**, 953–69.

Stelle, K. S. (1978). Classical gravity with higher derivatives. *Gen. Relativ. Gravit.*, **9**, 353–71.

Stoney, G. J. (1881). On the physical units of nature. *Philos. Mag., Ser. 5*, **11**, 381–90.

Straumann, N. (2000). Reflections on gravity. arXiv:astro-ph/0006423v1 [26 pp.].

Straumann, N. (2004). *General relativity with applications to astrophysics.* Springer, Berlin.

Strominger, A. (2009). Five problems in quantum gravity. *Nucl. Phys. Proc. Suppl.*, **192–193**, 119–25.

Strominger, A. and Vafa, C. (1996). Microscopic origin of the Bekenstein–Hawking entropy. *Phys. Lett. B*, **379**, 99–104.

Strumia, A. (2011). The fine-tuning price of the early LHC. *J. High Energy Phys.*, **04**, 073 [7 pp.].

Stueckelberg, E. C. G. (1938). Die Wechselwirkungskräfte in der Elektrodynamik und in der Feldtheorie der Kernkräfte (Teil 1). *Helv. Phys. Acta*, **11**, 225–44.

Sundermeyer, K. (1982). *Constrained dynamics.* Lecture Notes in Physics 169. Springer, Berlin.

Susskind, L. (2003). The anthropic landscape of string theory. arXiv:hep-th/0302219v1 [21 pp.].

Symanzik, K. (1981). Schrödinger representation and Casimir effect in renormalizable quantum field theory. *Nucl. Phys. B*, **190** [FS3], 1–44.

Tabensky, R. and Teitelboim, C. (1977). The square root of general relativity. *Phys. Lett. B*, **69**, 453–6.

Taylor, W. (2009). String field theory. In *Approaches to quantum gravity* (ed. D. Oriti), pp. 210–28. Cambridge University Press, Cambridge.

Tegmark, M., Aguirre, A., Rees, M. J., and Wilczek, F. (2006). Dimensionless constants, cosmology, and other dark matters. *Phys. Rev. D*, **73**, 023505 [28 pp.].

Teitelboim, C. (1977). Supergravity and square roots of constraints. *Phys. Rev. Lett.*, **38**, 1106–10.

Teitelboim, C. (1980). The Hamiltonian structure of space-time. In *General relativity and gravitation*, Vol. 1 (ed. A. Held), pp. 195–225. Plenum, New York.

Teitelboim, C. (1982). Quantum mechanics of the gravitational field. *Phys. Rev. D*, **25**, 3159–79.

Teitelboim, C. (1984). The Hamiltonian structure of two-dimensional space–time. In *Quantum theory of gravity* (ed. S. M. Christensen), pp. 327–44. Adam Hilger, Bristol.

Teitelboim, C. and Zanelli, J. (1987). Dimensionally continued topological gravitation theory in Hamiltonian form. *Class. Quantum Grav.*, **4**, L125–9.

Thiemann, T. (1996). Anomaly-free formulation of non-perturbative, four-dimensional, Lorentzian quantum gravity. *Phys. Lett. B*, **380**, 257–64.

Thiemann, T. (2006). The Phoenix project: master constraint programme for loop quantum gravity. *Class. Quantum Grav.*, **23**, 2211–48.

Thiemann, T. (2007). *Modern canonical quantum general relativity.* Cambridge University Press, Cambridge.

Thiemann, T. and Kastrup, H. A. (1993). Canonical quantization of spherically symmetric gravity in Ashtekar's self-dual representation. *Nucl. Phys. B*, **399**, 211–58.

Thirolf, P. G., Habs, D., Henig, A., Jung, D., Kiefer, D., Lang, C., Schreiber, J., Maia, C., Schaller, G., Schützhold, R., and Tajima, T. (2009). Signatures of the Unruh effect via high-power, short-pulse lasers. *Eur. Phys. J. D*, **55**, 379–89.

't Hooft, G. and Veltman, M. (1972). Combinatorics of gauge fields. *Nucl. Phys. B*, **50**, 318–53.

't Hooft, G. and Veltman, M. (1974). One-loop divergencies in the theory of gravitation. *Ann. Inst. Henri Poincaré A*, **20**, 69–94.

Tinkham, M. (2004). *Introduction to superconductivity.* Dover Publications, Mineola.

Toms, D. J. (2010). Quantum gravitational contributions to quantum electrodynamics. *Nature*, **468**, 56–9.

Torre, C. G. (1993*a*). Is general relativity an 'already parametrized' theory? *Phys. Rev. D*, **46**, 3231–4.

Torre, C. G. (1993*b*). Gravitational observables and local symmetries. *Phys. Rev. D*, **48**, R2373–6.

Torre, C. G. (1999). Midisuperspace models of canonical quantum gravity. *Int. J. Theor. Phys.*, **38**, 1081–102.

Torre, C. G. and Varadarajan, M. (1999). Functional evolution of free quantum fields. *Class. Quantum Grav.*, **16**, 2651–68.

Trautman, A. (1962). Conservation laws in general relativity. In *Gravitation: an introduction to current research* (ed. L. Witten), pp. 169–98. Wiley, New York.

Tryon, E. P. (1973). Is the universe a vacuum fluctuation? *Nature*, **246**, 396–7.

Tsamis, N. C. and Woodard, R. P. (1987). The factor-ordering problem must be regulated. *Phys. Rev. D*, **36**, 3641–50.

Tsamis, N. C. and Woodard, R. P. (1993). Relaxing the cosmological constant. *Phys. Lett. B*, **301**, 351–7.

Unruh, W. G. (1976). Notes on black-hole evaporation. *Phys. Rev. D*, **14**, 870–92.

Unruh, W. G. (1984). Steps towards a quantum theory of gravity. In *Quantum theory of gravity* (ed. S. M. Christensen), pp. 234–42. Adam Hilger, Bristol.

Uzan, J.-P. (2003). The fundamental constants and their variation: observational status and theoretical motivations. *Rev. Mod. Phys.*, **75**, 403–55.

van Dam, H. and Veltman, M. J. (1970). Massive and massless Yang–Mills and gravitational fields. *Nucl. Phys. B*, **22**, 397–411.

van de Ven, A. E. M. (1992). Two-loop quantum gravity. *Nucl. Phys. B*, **378**, 309–66.

van Nieuwenhuizen, P. (1981). Supergravity. *Phys. Rep.*, **68**, 189–398.

Vanzo, L., Acquaviva, G., and Di Criscienzo, R. (2011). Tunnelling methods and Hawkings radiation: achievements and prospects. arXiv:1106.4153v1 [gr-qc] [101 pp.].

van Zoest, T. and 27 other authors (2010). Bose–Einstein condensation in microgravity. *Science*, **328**, 1540–3.

Vaz, C., Witten, L., and Singh, T. P. (2001). Towards a midisuperspace quantization of LeMaître–Tolman–Bondi collapse models. *Phys. Rev. D*, **63**, 104020 [10 pp.].

Vaz, C., Gutti, S., Kiefer, C., Singh, T. P., and Wijewardhana, L. C. R. (2008*a*). Mass spectrum and statistical entropy of the BTZ black hole from canonical quantum gravity. *Phys. Rev. D*, **77**, 064021 [9 pp.].

Vaz, C., Tibrewala, R., and Singh, T. P. (2008*b*). Classical and quantum gravitational collapse in d-dimensional AdS spacetime. II. Quantum states and Hawking radiation. *Phys. Rev. D*, **78**, 024019 [12 pp.].

Veneziano, G. (1993). Classical and quantum gravity from string theory. In *Classical and quantum gravity* (ed. M. C. Bento, O. Bertolami, J. M. Mourão, and R. F. Picken), pp. 134–80. World Scientific, Singapore.

Vilenkin, A. (1982). Creation of universes from nothing. *Phys. Lett. B*, **117**, 25–8.

Vilenkin, A. (1988). Quantum cosmology and the initial state of the Universe. *Phys. Rev. D*, **37**, 888–97.

Vilenkin, A. (1989). Interpretation of the wave function of the Universe. *Phys. Rev. D*, **39**, 1116–22.

Vilenkin, A. (2003). Quantum cosmology and eternal inflation. In *The future of theoretical physics and cosmology* (ed. G. W. Gibbons, E. P. S. Shellard, and S. J. Rankin), pp. 649–66. Cambridge University Press, Cambridge.

Vilkovisky, G. (1984). The Gospel according to DeWitt. In *Quantum theory of gravity* (ed. S. M. Christensen), pp. 169–209. Adam Hilger, Bristol.

Visser, M. (2003). The quantum physics of chronology protection. In *The future of theoretical physics and cosmology* (ed. G. W. Gibbons, E. P. S. Shellard, and S. J. Rankin), pp. 161–76. Cambridge University Press, Cambridge.

von Borzeszkowski, H.-H. and Treder, H.-J. (1988). *The meaning of quantum gravity.* Reidel, Dordrecht.

von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik.* Springer, Berlin. For an English translation of parts of this book, see Wheeler, J. A. and Zurek, W. H. (eds.), *Quantum theory and measurement.* Princeton University Press, Princeton, NJ (1983).

von Weizsäcker, C. F. (1939). Der Zweite Hauptsatz und der Unterschied zwischen Vergangenheit und Zukunft. *Ann. Phys. (Berlin)*, **428**, 275–83.

Wald, R. M. (1984). *General relativity.* University of Chicago Press, Chicago.

Wald, R. M. (2001). The thermodynamics of black holes. *Living Rev. Relativ.*, **4**, article number 6.

Walls, D. F. and Milburn, G. J. (1994). *Quantum optics.* Springer, Berlin.

Weinberg, S. (1964). Photons and gravitons in S-matrix theory: derivation of charge conservation and equality of gravitational and inertial mass. *Phys. Rev.*, **135**, B1049–56.

Weinberg, S. (1972). *Gravitation and cosmology: principles and applications of the general theory of relativity.* Wiley, New York.

Weinberg, S. (1977). Critical phenomena for field theorists. In *Understanding the fundamental constituents of matter* (ed. A. Zichichi), pp. 1–52. Plenum Press, New York.

Weinberg, S. (1979). Ultraviolet divergences in quantum gravity. In *General relativity* (ed. S. W. Hawking and W. Israel), pp. 790–831. Cambridge University Press, Cambridge.

Weinberg, S. (1993). *Dreams of a final theory.* Hutchinson Radius, London.

Weinberg, S. (1995). *The quantum theory of fields, Vol. I (Foundations).* Cambridge University Press, Cambridge.

Weinberg, S. (1996). *The quantum theory of fields, Vol. II (Modern applications).* Cambridge University Press, Cambridge.

Weinberg, S. (1997). What is quantum field theory, and what did we think it is? arXiv:hep-th/9702027v1 [17 pp.].

Weinberg, S. (2000). *The quantum theory of fields, Vol. III (Supersymmetry).* Cambridge University Press, Cambridge.

Weinberg, S. (2007). Living in the multiverse. In *Universe or multiverse?* (ed. B. J. Carr), pp. 29–42. Cambridge University Press, Cambridge.

Weiss, P. (1938). On the Hamilton-Jacobi theory and quantization of generalized electrodynamics. *Proc. R. Soc. A*, **169**, 119–33.

Werner, S. A. and Kaiser, H. (1990). Neutron interferometry—macroscopic manifestations of quantum mechanics. In *Quantum mechanics in curved space–time* (ed. J. Audretsch and V. de Sabbata), pp. 1–21. Plenum Press, New York.

Wess, J. and Bagger, J. (1992). *Supersymmetry and supergravity*, 2nd edn. Princeton University Press, Princeton, NJ.

Wetterich, C. (1993). Exact evolution equation for the effective potential. *Phys. Lett. B*, **301**, 90–4.

Wetterich, C. (2011). Lattice spinor gravity. *Phys. Lett. B*, **704**, 612–9.

Wheeler, J. A. (1955). Geons. *Phys. Rev.*, **97**, 511–36.

Wheeler, J. A. (1968). Superspace and the nature of quantum geometrodynamics. In *Battelle rencontres* (ed. C. M. DeWitt and J. A. Wheeler), pp. 242–307. Benjamin, New York.

Wheeler, J. A. (1990). Information, physics, quantum: the search for links. In *Complexity, entropy, and the physics of information* (ed. W. H. Zurek), pp. 3–28. Addison-Wesley, Redwood City.

Wiltshire, D. L. (1996). An introduction to quantum cosmology. In *Cosmology: the physics of the Universe* (ed. B. Robson, N. Visvanathan, and W. S. Woolcock), pp. 473–531. World Scientific, Singapore. For a related version, see arXiv:gr-qc/-0101003v2 [60 pp.].

Witten, E. (1986). Non-commutative geometry and string field theory. *Nucl. Phys. B*, **268**, 253–94.

Witten, E. (1988). $2+1$ dimensional gravity as an exactly soluble system. *Nucl. Phys. B*, **311**, 46–78.

Witten, E. (1995). String theory in various dimensions. *Nucl. Phys. B*, **443**, 85–126.

Witten, E. (2003). A note on the Chern–Simons and Kodama wavefunctions. arXiv:gr-qc/0306083v2 [10 pp.].

Witten, E. (2007). Three-dimensional gravity reconsidered. arXiv:0706.3359v1 [hep-th] [83 pp.].

Woodard, R. P. (1993). Enforcing the Wheeler–DeWitt constraint the easy way. *Class. Quantum Grav.*, **10**, 483–96.

Woodard, R. P. (2009). How far are we from the quantum theory of gravity? *Rep. Prog. Phys.*, **72**, 126002 [42 pp].

Woodhouse, N. M. J. (1992). *Geometric quantization*, 2nd edn. Clarendon Press, Oxford.

Yoneya, T. (1974). Connection of dual models to electrodynamics and gravidynamics. *Progr. Theor. Phys.*, **51**, 1907–20.

Zakharov, V. I. (1970). Linearized gravitation theory and the graviton mass. *JETP Lett.*, **12**, 312–4.

Zeh, H. D. (1970). On the interpretation of measurement in quantum theory. *Found. Phys.*, **1**, 69–76. Reprinted in Wheeler, J. A. and Zurek, W. H. (eds.), *Quantum theory and measurement*. Princeton University Press, Princeton, NJ (1983).

Zeh, H. D. (1986). Emergence of classical time from a universal wave function. *Phys. Lett. A*, **116**, 9–12.

Zeh, H. D. (1988). Time in quantum gravity. *Phys. Lett. A*, **126**, 311–7.

Zeh, H. D. (1994). Decoherence and measurements. In *Stochastic evolution of quantum states in open systems and measurement processes* (ed. L. Diósi and B. Lukács), pp. 128–41. World Scientific, Singapore.

Zeh, H. D. (2003). There is no 'first' quantization. *Phys. Lett. A*, **309**, 329–34.

Zeh, H. D. (2005). Where has all the information gone? *Phys. Lett. A*, **347**, 1–7.

Zeh, H. D. (2007). *The physical basis of the direction of time*, 5th edn. Springer, Berlin.

Zeh, H. D. (2011). Feynman's quantum theory. *Eur. Phys. J. H*, **36**, 147–58.

Zel'dovich, Ya. B. and Starobinsky, A. A. (1984). Quantum creation of a universe with nontrivial topology. *Sov. Astron. Lett.*, **10**, 135–7.

Zhuk, A. (1992). Integrable multidimensional quantum cosmology. *Class. Quantum Grav.*, **9**, 2029–38.

Zurek, W. H. (2003). Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.*, **75**, 715–75.

Zwiebach, B. (2009). *A first course in string theory*, 2nd edn. Cambridge University Press, Cambridge.

Index

absolute elements, 78
accelerated observer, 211
acoustic horizon, 219
ADD scenario, 320
ADM action, 119
ADM energy, 128
AdS/CFT correspondence, 326
affine quantum gravity, 142
algebra of surface deformations, 95
anomalies, 161
anthropic principle, 352
anti-Chaplygin gas, 295
area law for black holes, 209
area metric, 202
area operator, 201
area spectrum
  for black holes, 228
  from loop quantum gravity, 201
arrow of time, 343
Ashtekar variables, 133, 135
asymptotic safety, 66
axial anomaly, 162
axion, 303

back reaction, 174, 217
  in semiclassical gravity, 178
background-field method, 47
background independence, 26, 47, 69, 78
  in string theory, 328
background manifold, 234
Baierlein–Sharp–Wheeler action, 124
Barbero–Immirzi parameter, 135, 230
Barbour–Bertotti model, 96
BCS state, 213
beables, 143
Bekenstein–Hawking entropy, 219
  from loop quantum gravity, 230
  from LTB model, 249
  from string theory, 317
  logarithmic corrections, 232
  statistical foundation, 219
Bergmann–Komar group, 126, 234
Berry phase, 174
BFV quantization, 52
Bianchi identities, 101
big brake, 295
big freeze, 296
big rip, 272, 295
billiards, cosmological, 296
BKL oscillations, 3, 348

black hole
  formation in accelerators, 324
  primordial, 252
  remnant, 250
black-hole entropy, *see* Bekenstein–Hawking entropy
black-hole spectroscopy, 228
black-hole thermodynamics
  laws of, 209
Bohr compactification, 291
Bohr–Rosenfeld analysis, 18
Boltzmann brain, 346
Born–Oppenheimer approximation, 172
  for Wheeler–DeWitt equation, 177
Bose–Einstein condensates in free fall, 10
bounce in loop quantum cosmology, 293, 297
BPS states, 317
brane worlds, 319
Brans–Dicke parameter, 308
BRST quantization, 52, 304
BTZ black hole, 274
bubble time, 147
Bunch–Davies vacuum, 283

Calabi–Yau manifold, 309
Callan–Symanzik equation, 67
canonical gauge, 83
canonical quantization, 141
causal sets, 26, 207
central charge, 89
central term, 89, 161
CGHS model, 163, 222
Chan–Paton factors, 313
chaotic inflation, 265
Chaplygin gas, generalized, 296
Chern–Simons action, 192
chiral perturbation theory, 57
Codazzi equation, 117
compactification, 309
condensed notation, 59
conformal factor problem, 44
conformal field theories, 88
conformal transformations, 88
connection dynamics, 133
constrained systems, 100
constraint algebra, 94
constraints
  first class, 103
  primary, 103, 120
  second class, 103, 168

secondary, 103
Copenhagen interpretation, 21
cosmic censorship, 209
cosmological constant, 58, 72, 352
covariant gauge fixing, 234
CPT theorem, 330
critical dimension in string theory, 302
critical exponents, 68
cylindrical function, 194

D-branes, 310, 313
dark energy, 70, 272
dark matter, 70
Davies–Unruh temperature, 15, 213
decoherence, 251
in quantum cosmology, 334
in quantum mechanics, 332
of primordial fluctuations, 342
deformed special relativity, 330
degrees of freedom, number of, 122
density matrix, reduced, 333
deparametrization, 80, 93, 146
DeWitt coefficients, 64
DeWitt metric, 111, 129
diffeomorphism constraints, 91, 121
quantum, 149, 152
dilaton, 163, 303, 304
dilaton gravity, 222
dimension
Hausdorff, 73
of space–time, 73
spectral, 73
dimensional regularization, 54
Dirac brackets, 168
Dirac conjecture, 104
Dirac consistency, 161
Dirac equation, 10
Dirac quantization, 81
Dirac–Bergmann algorithm, 102
Dirichlet condition in string theory, 299, 313
dust action, 243
dust shell
collapse, 233
quantization, 238
dynamical triangulation, 45, 70

effective action, 60, 307
effective average action, 68
effective theory, 26, 46
Einstein frame, 308
Einstein's equations, 2
linearized, 28
semiclassical, 15, 59, 63, 178
Einstein–Cartan theory, 2, 75
Einstein–Hilbert action, 1
Einstein–Hilbert truncation, 68
Einstein–Schrödinger equation, 149
embedding variables, 146
for dust shells, 235
emergent gravity, 25

energy–momentum tensor, 2
entanglement entropy, 251
ephemeris time, 99
equivalence principle, 12, 47
quantum tests, 10
Euclidean vacuum, 283, 336
Euler–Heisenberg Lagrangian, 47
Everett interpretation, 334
evolving constants, 126
extremal black holes
entropy, 227
quantization, 226
extrinsic curvature, 116
extrinsic time, 148, 263

factor ordering, 93, 142, 149, 288
in quantum cosmology, 265
Faddeev–Popov determinant, 51
Faddeev–Popov ghosts, 51
Faddeev–Popov procedure, 50, 182
Feynman diagrams
for quantum-gravitational corrections, 185
for quantum gravity, 48
field theories, parametrized, 90
Fierz–Pauli Lagrangian, 31
Friedmann universe, quantization, 262

gauge fixing, 50, 51, 105
gauge invariance
for linearized gravity, 36
of electrodynamics, 37
gauge orbit, 50
gauge theories, 104
from constraint algebra, 111
Gauss constraint, 134, 136, 191
for charged black holes, 222
Gauss equation, 117
Gauss's law, 112
Gauss–Bonnet theorem, 54
Gaussian fixed point, 68
Gaussian wave functional, 41, 158
Gell-Mann–Low equation, 67
general covariance, 78
generating functional, 59
geometric quantization, 141
ghost loops, 54
Gibbons–Hawking entropy, 347
global-time problem, 147
globally hyperbolic, 114
graceful exit problem, 272
gravitational bremsstrahlung, 58
gravitational radiation for quantum systems, 23
gravitational waves, weak, 28
gravitino, 76, 166
graviton, 28, 36, 186
composite, 74
from string theory, 303
graviton emission in atoms, 39
graviton loop, 63

graviton propagator, 185
graviton–graviton scattering, 57, 309
Green function, connected, 59
greybody factor, 215
Gribov ambiguities, 50
Gross–Neveu model, 65
ground state
  of free QED, 41
  of linear quantum gravity, 41
group averaging procedure, 144, 197
group field theory, 26, 207
group quantization, 238
GSO projection, 315
Gupta–Bleuler method, 36

Haar measure, 194
Hagedorn temperature, 318
Hamilton–Jacobi equation, 155, 175
Hamiltonian
  Dirac, 103
  total, 103
Hamiltonian constraint, 91, 121
  in loop quantum gravity, 203
Hamiltonian formulation of GR, 106
harmonic condition, 28, 53
harmonic oscillator, indefinite, 163, 269
Hartle–Hawking proposal, 279
Hartle–Hawking vacuum, 218
Hawking radiation, tunnelling interpretation, 218
Hawking temperature, 14, 213
helicity, 30, 35
heterotic string, 316
hierarchy problem, 320
Higgs field as inflaton, 287
higher-derivative quantum gravity, 49
Hilbert space
  in connection dynamics, 193
  Klein–Gordon type, 150
  problem of, 144
  rigged, 144, 195
  Schrödinger type, 149
Hořava–Lifshits gravity, 49, 74, 130
holographic principle, 326
holonomy, 139
homogeneous Lagrangians, 79
horizon, 211
horizontal stacking, 97, 124
Hubble parameter, 263
hybrid dynamics, 24

inflationary universe, 3, 255, 341
information-loss problem, 220, 241, 249
inhomogeneities in quantum cosmology, 275
interconnection theorems, 118
intrinsic time, 266
Ising model, 73
isolated horizon, 230
it from bit, 229

Jacobi action, 99
Jeans length, 253
Jordan frame, 308

Kalb–Ramond field, 303
kappa-Minkowski space–time, 330
KK tower, 321
Klein–Gordon equation, 82, 176

label time, 79
lapse function, 80, 82, 91, 115
large diffeomorphisms, 153
Leibniz group, 97
Lemaître–Tolman–Bondi model, 241
Levi–Civita connection, 134
Lichnerowicz equation, 148
little group, 35
loop algebra, 140
loop quantum gravity, 190
loop variables, 139, 193
Lorentz invariance, violation, 330
Lorenz gauge, 29

M-theory, 316, 326
many-fingered time, 93
Mashhoon effect, 12
mass-shell condition, 81
master constraint programme, 207
matrix theory, 316
mean field, 59
measurement analysis, 16
midisuperspace, 221, 261
minimal length, 20
minisuperspace, 261
mixmaster universe, 262
moduli space, 273, 307
momentum constraints, 91, 121
  quantum, 149, 152
momentum modes, 311
Mondrian diagrams, 77
Monte Carlo simulation, 73
multiple-choice problem, 147, 267
multiverse, 346, 349

Nambu–Goto action, 85
Neumann condition in string theory, 299
neutron interferometry, 8
Neveu–Schwarz boundary conditions, 315
Newtonian potential, quantum-gravitational correction, 57
no-boundary proposal, 279, 285
no-hair conjecture, cosmological, 3
no-hair theorem
  for black holes, 14, 208
Noether's theorems, 101
non-canonical gauge, 84
non-commutative geometry, 275, 314, 330
non-metric phase, 307
non-relativistic particle, parametrized, 78
non-renormalizability of quantum gravity, 45

observables, 105, 125, 128, 143, 144
Olbers' paradox, 344
one-loop approximation, 53
one-loop divergences, 55
one-loop effective action, 61, 286
open spaces, canonical formalism for, 127

path integral
  Euclidean, 43
  in string theory, 304
path integral composition law, 42
path-integral quantization, 42, 158
Pauli term, 46
perennials, 126
perturbation theory, 45
phantom fields, 272, 295
phase transition, 67
phenomenology, quantum-gravity, 329
photon, 35
  from string theory, 302
PLANCK mission, 188
Planck potential, 288
Planck units, 5
Poincaré gauge theory, 2
Poincaré group, 34
pointer states, 332
polarization tensor, 29
Polyakov action, 86, 304
polymer representation, 291
prefactor equation, 176, 182
primordial black holes, 252
problem of time, 4, 144

quantum connection dynamics, 190
quantum cosmology, 261
  from string theory, 271
  loop quantum cosmology, 289
  supersymmetric, 271
quantum cryptography, 23
quantum geometrodynamics, 141
quantum geometry, 195
quantum gravity in (2+1) dimensions, 273
quantum loop dynamics, 190
quasi-normal modes, 231, 251

Ramond boundary conditions, 315
Randall–Sundrum scenario, 322
rapidity, 225
Rarita–Schwinger action, 76
Rarita–Schwinger field, 76
reduced phase space, 105
reduced quantization, 146, 267
reference fluid, 148
Regge action, 72
Regge calculus, 44, 70
relational systems, 96
relativistic particle, 81
renormalizability, 46
renormalization group equation, 67
reparametrization invariance, 80, 85
Riemann–Cartan geometry, 12
Rindler space–time, 211
Rindler vacuum, 213
RNS action, 314

S-duality, 317
s-knot, 197
Sagnac effect, 9
sandwich conjecture, 123
Schmidt decomposition, 213
Schrödinger equation, 81, 147, 175
  for multipoles on three-sphere, 278
  from quantum gravity, 179
  quantum-gravitational corrections, 182
Schrödinger picture, functional, 142, 156
Schrödinger–Newton equation, 23
Schwarzschild black hole, canonical quantization, 221
Schwarzschild radius, 209
Schwarzschild–de Sitter solution, 249
Schwinger term, 89, 108, 161, 164
Schwinger–DeWitt technique, 64
second fundamental form, 116
self-dual connection, 137
semiclassical approximation, 172
shell, thin charged, 65
shift vector, 91, 115
sigma model, non-linear, 65, 304
singularities, naked, 226
singularity avoidance
  for the Wheeler–DeWitt equation, 295
  in loop quantum cosmology, 296
  in shell collapse, 240
singularity theorems, 2
Sommerfeld radiation condition, 344
space–time problem, 147
spin connection, 134
spin foam models, 207
spin network, 194, 195
squeezed states, 213
stochastic gravity, 63
string coupling constant, 306
string field theory, 86, 327
string interactions, 305
string landscape, 308, 328
string theory, 26, 298
  bosonic, 84, 299
strong-coupling limit, 127, 155
structural realism, 350
super-Hamiltonian, 79
supergravity (SUGRA), 74
  canonical, 165
  extended, 76
  from constraint algebra, 113
  $N = 8$, 77
superspace, 128, 152
  in string theory, 314
superstrings, 314
supersymmetry (SUSY), 74
surface gravity, 209

symmetric criticality principle, 262
symmetric initial condition, 288

T-duality, 310
tachyon, 302
target space, 304
tetrads, 11
theorema egregium, 117
theta vacuum, 153
third quantization, 151, 207, 328
time after quantization, 149
time before quantization, 145
time gauge, 133
time–energy uncertainty relations, 12
Tomonaga–Schwinger equation, 147, 156
torsion, 2, 166
trace anomaly, 187
trans-Planckian problem, 218
transverse-traceless gauge, 29
triads, 133
tunnelling proposal, 284, 285
twistors, 133
two-loop divergences, 55

uncertainty relation
  generalized, 310
  in string theory, 310
unification, 3
Unruh effect, 213
Unruh temperature, *see* Davies–Unruh temperature
Unruh vacuum, 213

vacua in string theory, 308
Vainshtein mechanism, 34
van Dam–Veltman–Zakharov discontinuity, 33
vertex operators, 306
vierbein, 11
Virasoro algebra, 88
volume operator, 204

wave packets in minisuperspace, 267
weak equality, 79, 103
Weyl anomaly, 64, 89, 162
Weyl invariance, 87
Weyl-tensor hypothesis, 347
Wheeler–DeWitt equation, 149, 177
  arrow of time from, 347
  for multipoles on three-sphere, 276
  in minisuperspace, 266
Wick rotation, 43
Wilson loop, 140
winding modes, 311
WKB approximation, 155, 172
WKB time, 175, 278, 341
WMAP satellite, 188

Yang–Mills theory, 113
York's time, 148, 274

zeta regularization, 286