Lecture Notes in Physics

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Physics and Astronomy ONLINE LIBRARY

http://www.springer.de/phys/

The Editorial Policy for Edited Volumes

The series *Lecture Notes in Physics* (LNP), founded in 1969, reports new developments in physics research and teaching - quickly, informally but with a high degree of quality. Manuscripts to be considered for publication are topical volumes consisting of a limited number of contributions, carefully edited and closely related to each other. Each contribution should contain at least partly original and previously unpublished material, be written in a clear, pedagogical style and aimed at a broader readership, especially graduate students and nonspecialist researchers wishing to familiarize themselves with the topic concerned. For this reason, traditional proceedings cannot be considered for this series though volumes to appear in this series are often based on material presented at conferences, workshops and schools.

Acceptance

A project can only be accepted tentatively for publication, by both the editorial board and the publisher, following thorough examination of the material submitted. The book proposal sent to the publisher should consist at least of a preliminary table of contents outlining the structure of the book together with abstracts of all contributions to be included. Final acceptance is issued by the series editor in charge, in consultation with the publisher, only after receiving the complete manuscript. Final acceptance, possibly requiring minor corrections, usually follows the tentative acceptance unless the final manuscript differs significantly from expectations (project outline). In particular, the series editors are entitled to reject individual contributions if they do not meet the high quality standards of this series. The final manuscript must be ready to print, and should include both an informative introduction and a sufficiently detailed subject index.

Contractual Aspects

Publication in LNP is free of charge. There is no formal contract, no royalties are paid, and no bulk orders are required, although special discounts are offered in this case. The volume editors receive jointly 30 free copies for their personal use and are entitled, as are the contributing authors, to purchase Springer books at a reduced rate. The publisher secures the copyright for each volume. As a rule, no reprints of individual contributions can be supplied.

Manuscript Submission

The manuscript in its final and approved version must be submitted in ready to print form. The corresponding electronic source files are also required for the production process, in particular the online version. Technical assistance in compiling the final manuscript can be provided by the publisher‘s production editor(s), especially with regard to the publisher’s own LaTeX macro package which has been specially designed for this series.

LNP Homepage (springerlink.com)

On the LNP homepage you will find:

—The LNP online archive. It contains the full texts (PDF) of all volumes published since 2000. Abstracts, table of contents and prefaces are accessible free of charge to everyone. Information about the availability of printed volumes can be obtained.

—The subscription information. The online archive is free of charge to all subscribers of the printed volumes.

—The editorial contacts, with respect to both scientific and technical matters.

—The author’s / editor’s instructions.

D. Giulini C. Kiefer C. Lämmerzahl (Eds.)

Quantum Gravity

From Theory to Experimental Search

Springer

Editors

Domenico J.W. Giulini
Universität Freiburg
Fakultät Physik und Mathematik
Hermann-Herder Str. 3
79104 Freiburg, Germany

Claus Lämmerzahl
Universität Bremen
ZARM
Am Fallturm
28359 Bremen, Germany

Claus Kiefer
Universität zu Köln
Institut Theoretischen Physik
Zülpicher Str. 77
50937 Köln, Germany

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at http://dnb.ddb.de

ISSN 0075-8450
ISBN 3-540-40810-X Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

springer.de

Printed in Germany

Typesetting: Camera-ready by the authors/editor
Cover design: *design & production*, Heidelberg

Printed on acid-free paper
54/3141/du - 5 4 3 2 1 0

Lecture Notes in Physics

For information about Vols. 1–586
please contact your bookseller or Springer-Verlag
LNP Online archive: springerlink.com/series/lnp/

Preface

The relation between quantum theory and the theory of gravitation is certainly one of the most outstanding unresolved issues of modern physics. On one side, quantum theory, in its usual formulation and orthodox interpretation, requires an ambient non-dynamical spacetime. On the other side, gravity, as described by general relativity, requires a dynamical geometry of spacetime which is coupled to all material processes within. This implies that at least one of these theories cannot be fundamentally correct. Hence, according to general expectation, there should exist a theory of quantum gravity comprising both previous theories. Such a theory should make definite predictions where previous theories failed to do so, like close to the Big Bang or during the radiational decay of Black Holes. Moreover, a theory of quantum gravity should also clarify the structure of spacetime at smallest scales. Up to now, no finally worked out theory of quantum gravity exists. Currently the most promising approaches to such a theory go under the names of Canonical Quantum Gravity and String Theory. The purpose of the 271st WE-Heraeus Seminar “Aspects of Quantum Gravity – From Theory to Experimental Search”, which took place in Bad Honnef from February 24th to March 1st, 2002, was to discuss issues surrounding quantum gravity on a level accessible to graduate students. The range of topics spanned an arc from fundamental questions concerning the notion of “quantisation”, over the presentation of definite approaches, to the possibility of astrophysical observations as well as laboratory experiments. We sincerely thank all speakers for their presentations and especially those who were moreover willing to write them up for the present volume. Last but not least we thank the Wilhelm and Else Heraeus Foundation for its generous support, without which this seminar could not have been realized, and the Physikzentrum for its kind hospitality.

Freiburg, Köln, and Bremen
April 2003

Domenico Giulini
Claus Kiefer
Claus Lämmerzahl

Contents

Overview

Part III Experimental Search

List of Contributors

Hartmut Abele
Physikalisches Institut
Universität Heidelberg
Philosophenweg 12
69120 Heidelberg
Germany
abele@physi.uni-heidelberg.de

Ignatios Antoniadis
CERN
Theory Division
1211 Geneva 23
Switzerland
Ignatios.Antoniadis@cern.ch

Stefan Baeßler
Physikalisches Institut
Universität Mainz
Staudingerweg 7
55099 Mainz
Germany
stefan.baessler@uni-mainz.de

Bernard J. Carr
Astronomy Unit
Queen Mary
University of London
Mile End Road
London E1 4NS
England
B.J.Carr@qmul.ac.uk

Domenico Giulini
Physikalisches Institut
Universität Freiburg
Hermann–Herder–Str. 3
79104 Freiburg
Germany
giulini@physik.uni-freiburg.de

Petr Hájíček
Institut für Theoretische Physik
Universität Bern
Sidlerstrasse 5
3012 Bern
Switzerland
hajicek@itp.unibe.ch

Claus Kiefer
Institut für Theoretische Physik
Universität zu Köln
Zülpicher Str. 77
50937 Köln
Germany
kiefer@thp.uni-koeln.de

Claus Lämmerzahl
ZARM
Universität Bremen
Am Fallturm
28359 Bremen
Germany
laemmerzahl@zarm.uni-bremen.de

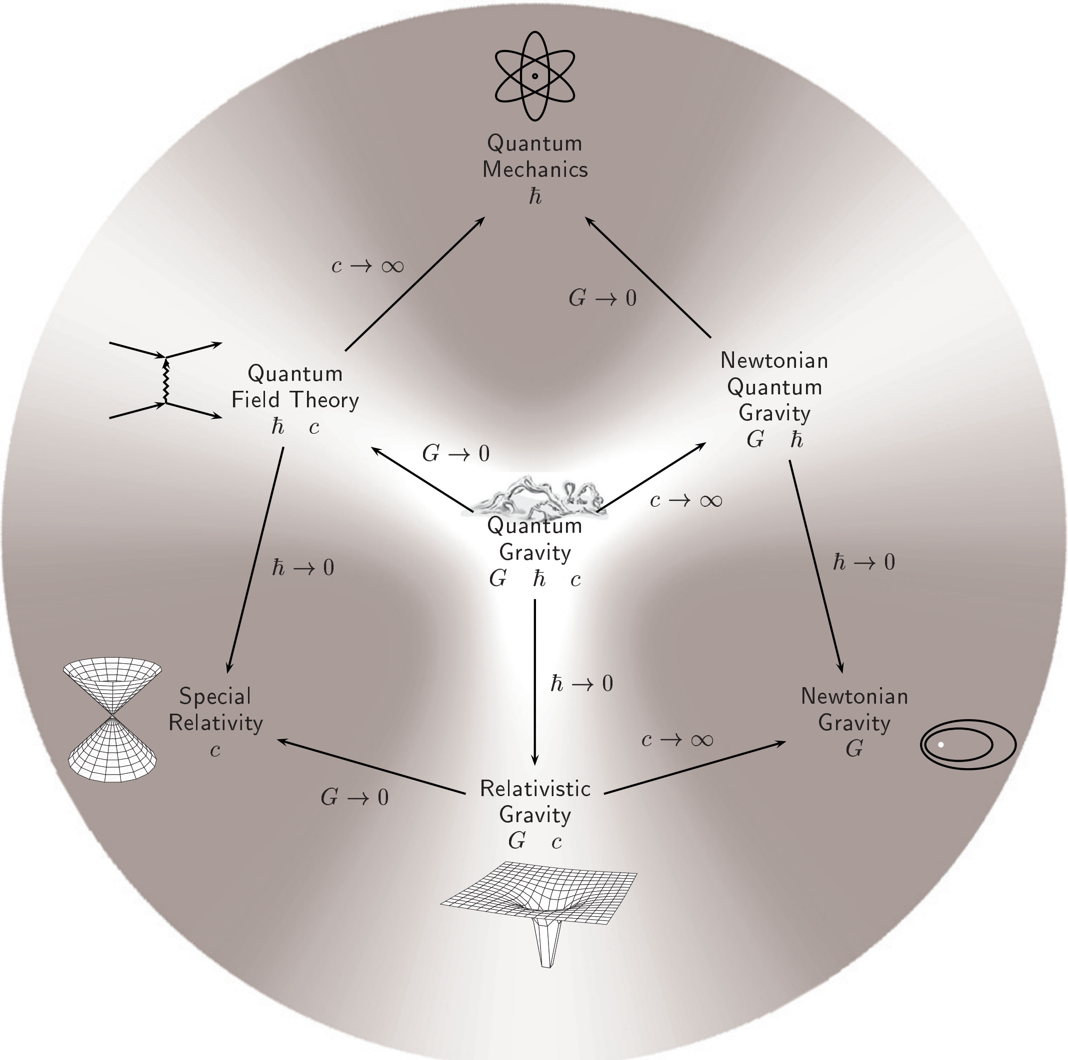
Renate Loll
Institute for Theoretical Physics
Utrecht University
Leuvenlaan 4
3584 CE Utrecht
The Netherlands
r.loll@phys.uu.nl

Thomas Mohaupt
Institute for Theoretical Physics
Friedrich–Schiller–Universität Jena
Max–Wien–Platz 1
07743 Jena
Germany
t.mohaupt@tpi.uni-jena.de

Daniel Sudarsky
Instituto de Ciencias Nucleares
Universidad Nacional Autónoma de México
A. Postal 70-543
México D.F. 04510
México
sudarsky@nuclecu.unam.mx

Thomas Thiemann
MPI für Gravitationsphysik
Albert-Einstein-Institut
Am Mühlenberg 1
14476 Golm
Germany
thiemann@aei-potsdam.mpg.de

Alexander Westphal
DESY Hamburg
Notkestraße 85
22607 Hamburg
Germany
westphal@mail.desy.de



Quantum
Mechanics
$\hbar$
$c \to \infty$
$G \to 0$
Quantum
Field Theory
$\hbar \quad c$
Newtonian
Quantum
Gravity
$G \quad \hbar$
$G \to 0$
$c \to \infty$
Quantum
Gravity
$G \quad \hbar \quad c$
$\hbar \to 0$
$\hbar \to 0$
$\hbar \to 0$
Special
Relativity
c
Newtonian
Gravity
G
$c \to \infty$
$G \to 0$
Relativistic
Gravity
$G \quad c$

Quantum Gravity – A General Introduction

Claus Kiefer

Institut für Theoretische Physik, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany.

Abstract. I give a brief introduction into the general problems of constructing a theory of quantum gravity, the main approaches, expected applications, as well as semiclassical approximations and the role of decoherence.

1 Quantum Theory and the Gravitational Field

Quantum theory seems to be a universal framework for physical theories. In fact, most of the interactions found in Nature are already successfully described by some quantum theory. The only interaction for which this has not yet been achieved is gravity. All manifestations of the gravitational field known so far can be understood from a classical theory – Einstein's theory of general relativity (also called 'geometrodynamics'). It is given by the Einstein-Hilbert action

$$S_{\mathrm{EH}} = \frac{c^4}{16\pi G} \int_{\mathcal{M}} \mathrm{d}^4x \, \sqrt{-g} \, (R - 2\Lambda) \; + \; \text{boundary term} + S_{\mathrm{m}} \; , \tag{1}$$

where S_{m} denotes the action for non-gravitational fields from which one can derive the energy-momentum tensor according to

$$T_{\mu\nu}(x) = \frac{2}{\sqrt{-g}} \frac{\delta S_{\mathrm{m}}}{\delta g^{\mu\nu}(x)} \; . \tag{2}$$

There exist certain 'uniqueness theorems' which state that every reasonable theory of the gravitational field must contain general relativity (or its natural generalisation, the Einstein-Cartan theory) in a certain limit, see e.g. [1] for a review.

In spite of its success, there are many reasons to believe that the most fundamental theory of gravity is a *quantum* theory. Unfortunately, no experimental material is presently available, which would point in a definite direction. The reasons are therefore of a theoretical nature. The main motivations for quantum gravity are [1]:

- **Unification.** The history of science shows that a reductionist viewpoint has been very fruitful in physics. The standard model of particle physics is a *quantum* field theory which has united in a certain sense all non-gravitational interactions. 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

C. Kiefer, Quantum Gravity – A General Introduction, Lect. Notes Phys. **631**, 3–13 (2003)
http://www.springerlink.com/

to now. This demonstrates in particular that classical and quantum *concepts* (phase space versus Hilbert space, etc.) are most likely incompatible.

- **Cosmology and Black Holes.** As the *singularity theorems* and the ensuing breakdown of general relativity 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 analogue of quantum mechanics (which due to its stationary states has rescued the atoms from collapse) the general expectation is that this encompassing theory is a *quantum* theory. It must be emphasised that *if* gravity is quantised, the kinematical nonseparability 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 below.
- **Problem of Time.** Quantum theory and general relativity (in fact, every general covariant theory) contain drastically different concepts of time (and spacetime). 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 spacetime). In contrast, spacetime is a dynamical object in general relativity. It is clear that a unification with quantum theory must lead to modifications of the concept of time. 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'.

What are the relevant scales on which effects of quantum gravity should be unavoidable? As has already been shown by Max Planck in 1899, the fundamental constants speed of light (c), gravitational constant (G), and quantum of action ($\hbar$) can be combined in a unique way (up to a dimensionless factor) to yield units of length, time, and mass. In Planck's honour they are called Planck length, $l_{\rm P}$, Planck time, $t_{\rm P}$, and Planck mass, $m_{\rm P}$, respectively. They are given by the expressions

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

$$t_{\rm P} = \frac{l_{\rm P}}{c} = \sqrt{\frac{\hbar G}{c^5}} \approx 5.40 \times 10^{-44}\ \text{s}\ , \tag{4}$$

$$m_{\rm P} = \frac{\hbar}{l_{\rm P} c} = \sqrt{\frac{\hbar c}{G}} \approx 2.17 \times 10^{-5}\ \text{g} \approx 1.22 \times 10^{19}\ \text{GeV}\ . \tag{5}$$

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 see direct quantum-gravity effects. In fact, the Planck scales are attained for an elementary particle whose Compton wavelength is (apart from a factor of 2) equal to its Schwarzschild radius,

$$\frac{\hbar}{m_{\mathrm{P}}c} \approx R_{\mathrm{S}} \equiv \frac{2Gm_{\mathrm{P}}}{c^2} \,, \tag{6}$$

which means that the spacetime curvature of an elementary particle is non-negligible. A truly unified theory may of course contain further parameters. An example is string theory (see next section) where the fundamental 'string length' l_{s} appears.

A quantity expressing the ratio of atomic scales to the Planck scale is the 'fine structure constant of gravity' defined by

$$\alpha_{\mathrm{g}} = \frac{Gm_{\mathrm{pr}}^2}{\hbar c} \equiv \left(\frac{m_{\mathrm{pr}}}{m_{\mathrm{P}}}\right)^2 \approx 5.91 \times 10^{-39} \,, \tag{7}$$

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 separation between micro- and macrophysics. It is interesting that structures in the universe occur for masses which can be expressed as simple powers of α_{g} in units of m_{pr}, cf. [2]. For example, stellar masses are of the order $\alpha_{\mathrm{g}}^{-3/2}m_{\mathrm{pr}}$, while stellar lifetimes are of the order $\alpha_{\mathrm{g}}^{-3/2}t_{\mathrm{P}}$. It is also interesting to note that the size of human beings is roughly the geometric mean of Planck length and size of the observable universe. It is an open question whether a fundamental theory of quantum gravity can provide an explanation for such values, e.g. for the ratio $m_{\mathrm{pr}}/m_{\mathrm{P}}$, or not. If not, only an anthropic principle could yield a – not very satisfying – 'explanation'.

Below the level of full quantum gravity one can distinguish from a conceptual point of view at least two other levels. The first, lowest, level deals with quantum *mechanics* in *external* gravitational fields (either described by general relativity 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, cf. the contribution by C. Lämmerzahl to this volume. Already in the 1970s, experiments of neutron interferometry were performed in the gravitational field of the Earth. It was possible, in particular, to show that the weak equivalence principle holds at the given level of precision. More recently, gravitational quantum bound states of neutrons in the field of the Earth have been measured, cf. the contribution by H. Abele.

The second level concerns quantum *field* theory in *external* gravitational fields. Back reaction can be taken into account in a perturbative sense. Although experimatal data are still lacking, there exist on this level at least precise predictions. The most important one concerns Hawking radiation for black holes [3], see e.g. [4] for a detailed review. A black hole radiates with temperature

$$T_{\mathrm{H}} = \frac{\hbar\kappa}{2\pi k_{\mathrm{B}}c} \,, \tag{8}$$

where κ is the surface gravity of a stationary black hole which by the no-hair theorem is uniquely characterised by its mass M, its angular momentum

J, and 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

$$T_{\rm H} = \frac{\hbar c^3}{8\pi k_{\rm B} GM} \approx 6.17 \times 10^{-8} \left(\frac{M_\odot}{M}\right) \text{ K} . \tag{9}$$

This temperature is unobservationally small for solar-mass (and bigger) black holes, but may be observable for primordial black holes, cf. the contribution by B. Carr. It must be emphasised that the expression for $T_{\rm H}$ contains all fundamental constants of nature. One may speculate that this expression – relating the macroscopic parameters of a black hole with thermodynamic quantities – plays a similar role for quantum gravity as de Broglie's relations $E = \hbar\omega$ and $p = \hbar k$ once played for the development of quantum theory [5]. Hawking radiation was derived in the semiclassical limit in which the gravitational field can be treated classically. According to (9), the black hole loses mass through its radiation and becomes hotter. After it has reached a mass of the size of the Planck mass (5), 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 (see below).

There exists a related effect to (8) in flat Minkowski space. An observer in uniform acceleration experiences the standard Minkowski vacuum not as empty, but as filled with *thermal* radiation with temperature

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

This temperature is often called the 'Davies-Unruh temperature', cf. [4]. Formally, it arises from (8) 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. Although (10) seems to be a small effect, it was suggested to search for it in accelerators or in experiments with ultra-intense lasers, without definite success up to now.

2 Approaches to Quantum Gravity

As I have already mentioned in the last section, experimental clues for quantum gravity are elusive. A direct probe of the Planck scale (5) in high-energy experiments would be illusory. In fact, an accelerator of current technology would have to have the size of several thousand lightyears in order to probe the Planck energy $m_{\rm P}c^2 \approx 10^{19}$ GeV. However, it is imaginable that effects of quantum gravity can in principle occur at lower energy scales. Possibilities could be non-trivial applications of the superposition principle for the quantised 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 spacetime 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 STEP, cf. the contribution by C. Lämmerzahl.

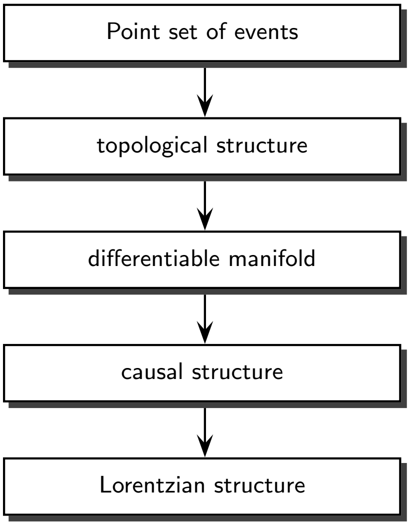
A truly fundamental theory should have such a rigid structure that all phenomena in the low-energy regime, such as particle masses or coupling constants, could be predicted in an unambiguous way. As there is no direct experimental hint yet, most work in quantum gravity focuses on the attempt to construct 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 Chris Isham makes a distinction between a 'primary theory of quantum gravity' and a 'secondary theory' [6]. In the primary approach, one starts with a given classical theory and applies heuristic quantisation rules. This is the approach usually adopted, and it was successful, for example, in QED. In most cases, 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, where 'covariant' refers here to spacetime diffeomorphisms. The main advantage of both approaches is that the starting point is given – the classical theory. 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 starts with a fundamental quantum framework of all interactions and tries to derive (quantum) general relativity in certain limiting situations, e.g. through an energy expansion. The most important example here is string theory (M-theory). The main advantage is that the fundamental quantum theory automatically yields a unification. The main disadvantage is that the starting point is entirely speculative. The general meaning of 'quantisation' is discussed in the contribution by D. Giulini.

Even if quantum general relativity is superseded by a more fundamental theory such as string theory, 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 general relativity should be a viable framework (such as QED, which is also supposed to be only an effective theory). It should also be mentioned that string theory automatically implements many of the methods used in the primary approach, such as quantisation of constrained systems and covariant perturbation theory.

An important question in the heuristic quantisation of a given classical theory is which of the structures in the classical theory should be quantised, i.e. subjected to the superposition principle, and which should remain as classical (or absolute, non-dynamical) structures. Isham distinguishes the following hierarchy of structures [7]:



Most approaches subject the Lorentzian and the causal structure to quantisation, but keep the manifold structure fixed. This is, however, not the only possibility. It might be that even the topological structure is fundamentally quantised. According to the Copenhagen interpretation of quantum theory, all these structures would probably have to stay classical, because they are thought to be necessary ingredients for the measurement process. For the purpose of quantum gravity, such a viewpoint is, however, insufficient and probably inconsistent.

Canonical quantum gravity is described in the contribution by T. Thiemann. Depending on the choice of canonical variables one distinguishes between various sub-approaches: quantum geometrodynamics, quantum connection dynamics, and quantum loop dynamics. Its central equations are the quantum constraints[1]

$$\hat{\mathcal{H}}_\perp \Psi = 0 \ , \tag{11}$$

$$\hat{\mathcal{H}}_a \Psi = 0 \ , \tag{12}$$

where (11) is usually referred to as the 'Wheeler-DeWitt equation' and (12) as the 'momentum' or 'diffeomorphism constraints' ($a = 1, 2, 3$). The argument of the wave functional Ψ is the space of all *three*-dimensional metrics $h_{ab}(\mathbf{x})$. Equations (12) guarantee, however, that Ψ is invariant under infinitesimal diffeomorphisms. The real arena is thus the space of all three-geometries ('superspace').

There are many problems associated with (11) and (12). Especially interesting from a conceptual point of view is the absence of an external time parameter t ('problem of time'). The reason is the dynamical nature that time plays in general relativity: on the one hand, it cannot appear as a classical time parameter

[1] In some approaches additional gauge constraints occur, see Thiemann's contribution.

like in ordinary quantum theory; on the other hand, the uncertainty relation in gravity forbids the simultaneous specification of three-geometry and second fundamental form, so the concept of spacetime is completely lost in the quantum theory. This is fully analogous to the loss of particle trajectories in quantum mechanics, see e.g. [1,8] for a detailed discussion and references. An issue related to the problem of time is the 'problem of Hilbert space': it is not known which Hilbert space, if any, has to be used for the physical degrees of freedom in the full theory. One therefore treats the quantum constraints (11) and (12) often pragmatically as differential equations, with boundary conditions being imposed from physical reasoning.

Quantum general relativity does not necessarily have to be treated in a canonical approach. Alternative methods are the traditional background field method and path-integral quantisation [1]. In the former, a perturbation is performed around a four-dimensional background metric, and four-dimensional covariance with respect to this metric is preserved at each order of perturbation theory. The theory is perturbatively non-renormalisable, so it loses its predictive power at high energies. Nevertheless, it is viable as an *effective theory* at low energies (in the infrared limit). In this limit one can calculate, for example, quantum gravitational corrections to Newton's law [1]. Quite generally it is expected that possible observations of a fundamental theory of quantum gravity can be described on the level of effective actions, e.g. concerning searches for non-Newtonian gravity or the violation of the weak equivalence principle, cf. the contributions by I. Antoniadis and C. Lämmerzahl. The path-integral approach is described in the contribution by R. Loll.

String theory is described in the contribution by T. Mohaupt. In contrast to quantum general relativity, it automatically yields a unified quantum framework for all interactions. Until around 1996 most developments in string theory occurred on the perturbative level. One of the main outcomes was that gravity is inevitable. Other predictions are the occurrence of gauge invariance, supersymmetry, and the presence of higher dimensions. The theory is envisioned to be free of infinities.

More recently, the study of non-perturbative aspects has emerged. This is mostly triggered by the occurrence of D-branes (higher-dimensional objects on which open strings can end) and the discovery of dualities. They allow to relate the small-coupling regime of one version of string theory to the large-coupling regime of another version.

The history of quantum gravity starts with early perturbative attempts by Leon Rosenfeld in 1929. A brief overview of historical developments can be found in [9].

3 Quantum Black Holes and Quantum Cosmology

It is expected that two of the main applications of any theory of quantum gravity concerns black holes and cosmology. For black holes, the level of quantum field theory on a fixed background (Sect. 1) leads to the concept of Hawking radiation, see (8) and (9). Connected with this temperature is the occurrence of the 'Bekenstein-Hawking entropy'

$$S_{\mathrm{BH}} = \frac{k_{\mathrm{B}} A}{4 G \hbar} \ , \tag{13}$$

where A is the surface of the event horizon. The black-hole entropy (13) is much bigger than the entropy of a collapsing star. The entropy of the Sun, for example, is $S_\odot \approx 10^{57}$, but the entropy of a solar-mass black hole is $S_{\mathrm{BH}} \approx 10^{77}$, i.e. twenty orders of magnitudes larger (all entropies are measured in units of k_{B}). If all matter in the observable Universe were in a single gigantic black hole, its entropy would be $S_{\mathrm{BH}} \approx 10^{123}$. Black holes thus seem to be the most efficient objects for swallowing information.

Due to Hawking radiation, black holes have a finite lifetime. It is given by

$$\tau_{\mathrm{BH}} \approx \left(\frac{M_0}{m_{\mathrm{P}}}\right)^3 t_{\mathrm{P}} \approx 10^{65} \left(\frac{M_0}{M_\odot}\right)^3 \text{years} \ . \tag{14}$$

It has been speculated that after this time a black hole has evaporated completely and has left behind only thermal radiation. This would be independent of any initial state the black hole has started from. Since a thermal state contains least information, one would then be faced with the *information-loss problem*. This is, however, a contentious issue and many arguments have been put forward in favour of a unitary evolution for the black hole, see e.g. [10]. The final word on this issue will be said after the full theory of quantum theory is known. Such a theory should also provide a derivation of (13) by counting microscopic quantum states. Preliminary results have been achieved both within the canonical approach [11] and string theory [12], cf. the contributions by T. Mohaupt and D. Sudarsky. Quantum gravity should also provide a detailed understanding of the final evaporation process and settle the question whether the area of the event horizon is quantised and, if yes, what its spectrum is.

To get a grip on the fate of the classical singularity, one can discuss exact models of quantum gravitational collapse. This is done in the contribution by P. Hájíček. He considers a thin spherically-symmetric shell with zero rest mass that classically collapses into a black-hole singularity. One can, however, construct a unitary quantum theory in which this singularity is avoided. If the shell is described as a wave packet, the initially purely-collapsing packet turns near the horizon into a superposition of collapsing and expanding packet and guarantees that the wave function is zero at $r \to 0$. For late times the packet will be fully expanding.

If quantum theory is applied to the universe as a whole, one talks about *quantum cosmology*. Since the dominating interaction on large scales is gravity, this can be described only within a quantum theory of gravity. Models can be constructed in all existing approaches by making symmetry assumptions such as homogeneity and isotropy. To discuss just one example, let us consider a closed Friedmann universe with scale factor ('radius') $a \equiv e^\alpha$ containing a massive scalar field ϕ with mass m. In this case the Wheeler-DeWitt equation (11) can be written in suitable units for a wave function $\psi(a,\phi)$ – the 'wave function of the universe' – as

$$\hat{H}\psi \equiv \left(\hbar^2\frac{\partial^2}{\partial\alpha^2} - \hbar^2\frac{\partial^2}{\partial\phi^2} + m^2\phi^2 e^{6\alpha} - e^{4\alpha}\right)\psi(\alpha,\phi) = 0\ . \tag{15}$$

One recognises explicitly the hyperbolic nature ('wave nature') of this equation. The role of intrinsic time is played by α; this becomes evident if further degrees of freedom are added: they all come with the sign of the kinetic term for ϕ.

Since no external time parameter t is contained in (11), one cannot pose any initial conditions with respect to it. Instead, one can specify the wave function (and its derivative) – in the example (15) – at a fixed value of α. This is the natural boundary condition for a hyperbolic equation. It has drastic consequences if one wants to describe a universe that classically expands, reaches a maximum and recollapses again [5]. Both big bang and big crunch correspond to the same region in configuration space – the region of $\alpha \to -\infty$. They are thus intrinsically indistinguishable. The Wheeler-DeWitt equation connects larger scale factors with smaller scale factors, but not two ends of a classical trajectory. If one wants to mimick the classical trajectory by a 'recollapsing' wave packet, one has to include both the 'initial' and the 'final' wave packet into one initial condition with respect to α. If one of the two packets were lacking, one would not be able to recover the classical trajectory as an approximation.

There is another interesting feature in the case of recollapsing universes: it is in general not possible to construct from (15) a wave packet that follows as a narrow tube the classical trajectory [5]. Therefore, a semiclassical approximation is not valid all along the trajectory and quantum effects can play a role even far away from the Planck scale – e.g. at the turning point of the classical universe.

Quantum-cosmological models such as (15) can serve quite generally to discuss the role of boundary conditions (e.g. the 'no-boundary condition' or the 'tunneling condition') [8] or issues related to the problem of time. An interesting question, for example, concerns the origin of the inflationary universe in a theory of quantum gravity [13].

4 Semiclassical Approximation and Decoherence

In order to bridge the gap between quantum gravity and the limit of quantum theory in an external background, some kind of approximation scheme must be devised. This has been discussed in all approaches, and I want to sketch here only the procedure in quantum geometrodynamics, see [8] for more details and references.

One method involves a Born-Oppenheimer type of approximation with respect to the Planck mass $m_{\rm P}$. The situation is formally similar to molecular physics where the heavy nuclei move slowly, followed adiabatically by the light electrons. In situations where the relevant scales are much smaller than the Planck mass, the gravitational kinetic term can be neglected in a first approximation. One makes for solutions of (11) the ansatz

$$\Psi[h_{ab},\varphi] = e^{\mathrm{i}m_{\mathrm{P}}^2 S[h_{ab}]}\Phi[h_{ab},\varphi] \ , \tag{16}$$

where φ stands symbolically for non-gravitational fields. Inserting this into (11) and and making an expansion with respect to m_{P}, one finds that $S[h_{ab}]$ obeys the gravitational Hamilton-Jacobi equation. This is known to be equivalent to Einstein's field equations. In this sense the classical background spacetime emerges as an approximation (such as geometrical optics emerges as a limit from wave optics). One can now pick out one classical spacetime from the many classical solutions (spacetimes) that are described by $S[h_{ab}]$. The 'matter wave functional' $\Phi[h_{ab},\varphi]$ can then be evaluated on this particular spacetime described by $h_{ab}(\mathbf{x},t)$ and can therefore shortly be labelled $\Phi(t,\varphi]$. If other semiclassical variables are present (such as the homogeneous field ϕ in (15)), they are included in S. The time parameter t is defined from $S[h_{ab}]$ as parametrising the classical trajectory (spacetime) running orthogonally to $S[h_{ab}] = \text{const.}$ in the space of three-geometries. In the special case (15) of the Friedmann universe, t is defined by the scale factor $a(t)$ and the homogeneous scalar field $\phi(t)$. It can be shown from (11) that the time evolution of the state Φ,

$$\frac{\partial}{\partial t}\,\Phi(t,\varphi] = \int d^3x\,\dot{h}_{ab}(\mathbf{x},t)\,\frac{\delta}{\delta h_{ab}(\mathbf{x},t)}\Phi[h_{ab}(\mathbf{x},t),\varphi] \ . \tag{17}$$

is given by a functional Schrödinger equation in the external classical spacetime found from $S[h_{ab}]$,

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\,\Phi(t,\varphi] = \hat{H}^{\mathrm{mat}}\Phi(t,\varphi] \ , \tag{18}$$

where $\hat{H}^{\mathrm{mat}}$ is the matter field Hamiltonian in the Schrödinger picture, parametrically depending on (generally nonstatic) metric coefficients of the curved spacetime background. In this way, the Schrödinger equation for non-gravitational fields has been recovered from quantum gravity as an approximation. A derivation similar to the above can already be performed within ordinary quantum mechanics if one assumes that the total system is in a 'timeless' energy eigenstate. In fact, Neville Mott had already considered in 1931 a time-independent Schrödinger equation for a total system consisting of an α-particle and an atom. If the state of the α-particle can be described by a plane wave (corresponding in this case to high velocities), one can make an ansatz similar to (16) and derive a time-dependent Schrödinger equation for the atom alone, in which time is *defined* by the α-particle.

Higher orders in this Born-Oppenheimer scheme yield quantum-gravitational correction terms to the Schrödinger equation, which could leave an observational imprint e.g. in the anisotropy spectrum of the cosmic microwave background.

The ansatz (16) is already special, since it is a product of a pure phase part depending on gravity with a matter wave function. The i in the Schrödinger equation (18) has its origin in the choice of this phase. Can this be justified?

The answer is yes. A crucial role is hereby played by the process of *decoherence* [14]. This is the emergence of classical properties through the irreversible interaction of a quantum system with its environment. Information about possible interference effects in the system is delocalised into quantum correlations with the inaccessible degrees of freedom of the environment and is no longer available at the system itself. Formally, decoherence is described through the reduced density matrix of the system obtained by tracing out the irrelevant degrees of freedom. In the present context these irrelevant variables can be density fluctuations or gravitational waves. Detailed discussions show that states of the form (16) are most robust against environmental influence and that the variables contained in $S[h_{ab}]$ assume quasiclassical properties [5,8,14]. It is also possible along these lines to understand, at least in principle, the origin of the arrow of time in our universe from a simple boundary condition in quantum cosmology [5,15].

References

1. C. Kiefer: *Quantum gravity* (Oxford University Press, Oxford, to appear)
2. M. Rees: *Perspectives in astrophysical cosmology* (Cambridge University Press, Cambridge, 1995)
3. S.W. Hawking: Commun. Math. Phys. **43**, 199 (1975)
4. C. Kiefer: Thermodynamics of black holes and Hawking radiation. In: *Classical and quantum black holes*, ed. by P. Fré, V. Gorini, G. Magli, U. Moschella (IOP Publishing, Bristol, 1999)
5. H.D. Zeh: *The physical basis of the direction of time*, 4th edition (Springer, Berlin, 2001). See also `http://www.time-direction.de`
6. C.J. Isham: Quantum gravity. In: *General relativity and gravitation*, ed. by M.A.H. Mac Callum (Cambridge University Press, Cambridge, 1987)
7. C.J. Isham: Prima facie questions in quantum gravity. In: *Canonical gravity: From classical to quantum*, ed. by J. Ehlers and H. Friedrich (Springer, Berlin, 1994)
8. C. Kiefer: in: *Towards quantum gravity*, ed. by J. Kowalski-Glikman (Springer, Berlin, 2000)
9. C. Rovelli: Notes for a brief history of quantum gravity. `gr-qc/0006061`
10. C. Kiefer: Is there an information-loss problem for black holes? To appear in: *Quantum decoherence and entropy in complex systems*, ed. by H.-T. Elze (Springer, Berlin, 2003)
11. A. Ashtekar, J.C. Baez, and K. Krasnov: Adv. Theor. Math. Phys. **4**, 1 (2000)
12. G.T. Horowitz: Quantum states of black holes. In: *Black holes and relativistic stars*, ed. by R.M. Wald (The University of Chicago Press, Chicago 1998)
13. A.O. Barvinsky, A.Y. Kamenshchik, and C. Kiefer: Mod. Phys. Lett. A **14**, 1083 (1999)
14. E. Joos, H.D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I.-O. Stamatescu: *Decoherence and the appearance of a classical world in quantum theory*, 2nd edition (Springer, Berlin, 2003). See also `http://www.decoherence.de`
15. C. Kiefer: Arrow of time from timeless quantum gravity. To appear in: *Time and matter*, edited by I. Bigi and M. Fäßler (World Scientific, Singapore, 2003)

That Strange Procedure Called Quantisation

Domenico Giulini

Physikalisches Institut, Universität Freiburg, Hermann-Herder-Str. 3,
9104 Freiburg, Germany

Abstract. I discuss the notion of 'quantisation' á la Dirac (canonical quantisation) from a general perspective. It is well known that Dirac's quantisation rules cannot work in general. I present this classic no-go result, which is due to Groenewold and van Hove, with due emphasis on its hypotheses. Finally, I briefly discuss first-class constrained systems with emphasis on the global-geometric and algebraic apsects.

1 Introduction and Motivation

In my contribution I wish to concentrate on some fundamental issues concerning the notion of *quantisation*. Nothing of what I will say is new or surprising to the experts. My intention is rather a pedagogical one: to acquaint the non-experts with some of the basic structural results in quantisation theory, which I feel should be known to anybody who intends to 'quantise' something. A central result is the theorem of Groenewold and van Hove, which is primarily a no-go result, stating that the most straightforward axiomatisation of Dirac's informally presented 'canonical' quantisation rules runs into contradictions and therefore has to be relaxed. The constructive value of this theorem lies in the fact that its proof makes definite suggestions for such relaxations. This helps to sharpen ones expectations on the quantisation concept in general, which is particularly important for Quantum Gravity since here sources for direct physical input are rather scarce. Expectations on what Quantum Gravity will finally turn out to be are still diverse, though more precise pictures now definitely emerge within the individual approaches, as you will hopefully be convinced in the other lectures (see the lectures by Loll, Mohaupt, and Thiemann in this volume) so that reliable statements about similarities and differences on various points can now be made. The present contribution deliberately takes focus on a very particular and seemingly formal point, in order to exemplify in a controllable setting the care needed in formulating 'rules' for 'quantisation'. At the end I will also briefly consider constrained systems from a slightly more 'global' point of view. Two appendices provide some technical aspects.

How do you recognize *quantum* theories and what structural elements distinguish them from so-called classical ones? If someone laid down, in mathematical terms, a theory of 'something' before you, what features would you check in order to answer this question? Or would you rather maintain that this question does not make good sense to begin with? Strangely enough, even though quantum theories are not only known to be the most successful but also believed to

D. Giulini, That Strange Procedure Called Quantisation, Lect. Notes Phys. **631**, 17–40 (2003)
http://www.springerlink.com/

be the most fundamental theories of physics, there seems to be no unanimously accepted answer to any of these questions. So far a working hypothesis has been to *define* quantum theories as the results of some 'quantisation procedures' after their application to classical theories. One says that the classical theory (of 'something') 'gets quantised' and that the result is the quantum theory (of that 'something'). This is certainly the way we traditionally understand Quantum Mechanics and also a substantial part of Quantum Field Theory (for more discussion on this point, that also covers interesting technical issues, I recommend [12]). As an exception – to a certain degree – I would list Local Quantum Field Theory [10], which axiomatically starts with a general kinematical framework for Poincaré invariant quantum field theories without any a priori reference to classical theories. Although this can now be generalised to curved spacetimes, it does not seem possible to eliminate the need of some such fixed (i.e. non-dynamical) background. Hence this approach does not seem to be able to apply to background independent dynamical fields, like gravity.

The generally accepted quantisation procedures I have in mind here can be roughly divided into three groups, with various interrelations:

- Hilbert-space based methods, like the standard canonical quantisation programme,
- algebraic methods based on the notion on observables, like $\star$-product quantisation or C^*-algebra methods,
- path integral methods.

Given the success of Quantum Mechanics (QM) it was historically, and still is, more than justified to take it as paradigm for all other quantum theories (modulo extra technical inputs one needs to handle infinitely many degrees of freedom). Let us therefore take a look at QM and see how quantisation may, or may not, be understood. In doing this, I will exclusively focus on the traditional 'canonical' approaches to quantisation.

2 Canonical Quantisation

Historically the rules for 'canonical quantisation' where first spelled out by Dirac in his famous book on QM [3]. His followers sometimes bluntly restated these rules by the symbolic line,

$$\{\cdot\,,\,\cdot\} \mapsto \tfrac{-i}{\hbar}[\cdot\,,\cdot]\,, \tag{1}$$

which is to be read as follows: map each classical observable (function on phase space) f to an operator $\hat{f}$ in a Hilbert space $\mathcal{H}$ (typically $L^2(Q, d\mu)$, where Q is the classical configuration space and $d\mu$ the measure that derives from the Riemannian metric thereon defined by the kinetic energy) in such a way that the Poisson bracket of two observables is mapped to $-i/\hbar$ times the commutator of the corresponding operators, i.e, $\widehat{\{f_1, f_2\}} = \frac{-i}{\hbar}[\hat{f}_1, \hat{f}_2]$ (see e.g. [1], Sect. 5.4). This is also facetiously known as 'quantisation by hatting'. But actually Dirac was more careful; he wrote [3] (my emphasis; P.B. denotes 'Poisson Brackets')

> 'The strong analogy between quantum P.B. [i.e. commutators] and classical P.B. leads us to make the assumption that the quantum P.B., *or at any rate the simpler ones of them,* have the same values as the corresponding classical P.B.s.'
>
> PAUL DIRAC, 1930

Clearly these words demand a specific interpretation before they can be called a (well defined) quantisation programme.

2.1 The Classical Stage

Associated to a classical Hamiltonian dynamical system of n degrees of freedom is a $2n$-dimensional manifold, P, the space of states or 'phase space' (sometimes identified with the space of solutions to Hamilton's equations, if the latter pose a well defined initial-value problem). Usually – but not always – it comes equipped with a preferred set of $2n$ functions, (q^i, p_i), $i = 1 \cdots n$, called coordinates and momenta respectively. In addition, there is a differential-geometric structure on P, called *Poisson Bracket*, which gives a suitable subspace $\mathcal{F} \subseteq C^\infty(P)$ of the space of real-valued, infinitely differentiable functions the structure of a Lie algebra. See Appendix 1 for more information on the geometric structures of classical phase space and Appendix 2 for the general definition of a Lie algebra. Exactly what subspace is 'suitable' depends of the situation at hand and will be left open for the time being. In any case, the Poisson Bracket is a map

$$\{\cdot,\cdot\}: \ \mathcal{F} \times \mathcal{F} \to \mathcal{F}\,, \tag{2}$$

which satisfies the following conditions for all $f, g, h \in \mathcal{F}(P)$ and $\lambda \in \mathbb{R}$ (which make it precisely a real Lie algebra):

$$\{f,g\} = -\{g,f\} \qquad \text{antisymmetry}\,, \tag{3}$$

$$\{f,g+\lambda h\} = \{f,g\} + \lambda\{f,h\} \qquad \text{linearity}\,, \tag{4}$$

$$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0 \qquad \text{Jacobi identity}\,. \tag{5}$$

In the special coordinates (q^i, p_i) it takes the explicit form (cf. Appendix 1)

$$\{f,g\} := \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \right)\,. \tag{6}$$

Independently of the existence of a Poisson Bracket, the space $\mathcal{F}$ is a commutative and associative algebra under the operation of pointwise multiplication:

$$(f \cdot g)(x) := f(x)g(x)\,. \tag{7}$$

This means that the multiplication operation is also a map $\mathcal{F} \times \mathcal{F} \to \mathcal{F}$ (simply denoted by '$\cdot$') which satisfies the following conditions for all $f, g, h \in \mathcal{F}$ and $\lambda \in \mathbb{R}$:

$$f \cdot g = g \cdot f \qquad \text{commutativity}\,, \tag{8}$$

$$f \cdot (g + \lambda\, h) = f \cdot g + \lambda\, f \cdot h \qquad \text{linearity}\,, \tag{9}$$

$$f \cdot (g \cdot h) = (f \cdot g) \cdot h \qquad \text{associativity}\,. \tag{10}$$

The two structures are intertwined by the following condition, which expresses the fact that each map $D_f : \mathcal{F} \to \mathcal{F}$, $g \mapsto D_f(g) := \{f, g\}$, is a derivation of the associative algebra for each $f \in \mathcal{F}$:

$$\{f, g \cdot h\} = \{f, g\} \cdot h + g \cdot \{f, h\}\,. \tag{11}$$

The Jacobi identity now implies that ($\circ$ denotes composition) $D_f \circ D_g - D_g \circ D_f = D_{\{f,g\}}$.[1] Taken all this together this makes $\mathcal{F}$ into a Poisson algebra, whose abstract definition is as follows:

Definition 1. A *Poisson algebra* is a vector space V with two maps $V \times V \to V$, denoted by '$\{\,,\}$' and '$\cdot$', which turn V into a Lie algebra (defined by (3-5)) and a commutative and associative algebra (defined by (8-10)) respectively, such that (11) holds.

Simply writing the symbol $\mathcal{F}$ now becomes ambiguous since it does not indicate which of these different structures we wish to be implicitly understood. I shall use the convention to let '+' indicate the vector-space structure, $(+, \{\,,\})$ the Lie-algebra structure, $(+, \cdot)$ the associative structure, and $(+, \{\,,\}, \cdot)$ the Poisson structure. To avoid confusion I will then sometimes write:

$$\mathcal{F} \qquad \text{for the set}\,, \tag{12}$$

$$\mathcal{F}(+, \{\,,\}) \qquad \text{for the Lie algebra}\,, \tag{13}$$

$$\mathcal{F}(+, \cdot) \qquad \text{for the associative algebra}\,, \tag{14}$$

$$\mathcal{F}(+, \{\,,\}, \cdot) \qquad \text{for the Poisson algebra}\,, \tag{15}$$

formed by our subset of functions from $C^\infty(P)$. Sometimes I will indicate the subset of functions by a subscript on $\mathcal{F}$. For example, I will mostly restrict P to be $\mathbb{R}^{2n}$ with coordinates (q^i, p_i). It then makes sense to restrict to functions which are polynomials in these coordinates.[2] Then the following subspaces will turn out to be important in the sequel:

---

[1] This can be expressed by saying that the assignment $f \mapsto D_f$ is a Lie homomorphism from the Lie algebra $\mathcal{F}$ to the Lie algebra of derivations on $\mathcal{F}$. Note that the derivations form an associative algebra when multiplication is defined to be composition, and hence also a Lie algebra when the Lie product is defined to be the commutator.

[2] Recall that you need an affine structure on a space in order to give meaning to the term 'polynomial functions'.

$$\mathcal{F}_\infty \quad : \quad C^\infty\text{-functions,} \tag{16}$$
$$\mathcal{F}_{\text{pol}} \quad : \quad \text{polynomials in } q\text{'s and } p\text{'s,} \tag{17}$$
$$\mathcal{F}_{\text{pol}(1)} \quad : \quad \text{polynomials of at most first order,} \tag{18}$$
$$\mathcal{F}_{\text{pol}(2)} \quad : \quad \text{polynomials of at most second order,} \tag{19}$$
$$\mathcal{F}_{\text{pol}(\infty,1)} : \quad \text{polynomials of at most first order in the } p\text{'s whose coefficients are polynomials in the } q\text{'s.} \tag{20}$$

An otherwise unrestricted polynomial dependence is clearly preserved under addition, scalar multiplication, multiplication of functions, and also taking the Poisson Bracket (6). Hence $\mathcal{F}_{\text{pol}}$ forms a Poisson subalgebra. This is not true for the other subspaces listed above, which still form Lie subalgebras but not associative algebras.

2.2 Defining 'Canonical Quantisation'

Roughly speaking, Dirac's approach to quantisation consists in mapping certain functions on P to the set SYM($\mathcal{H}$) of symmetric operators (sometimes called 'formally self adjoint') on a Hilbert space $\mathcal{H}$. Suppose these operators have a common invariant dense domain $\mathcal{D} \subset \mathcal{H}$ (typically the 'Schwarz space'), then it makes sense to freely multiply them. This generates an associative algebra of operators (which clearly now also contains non-symmetric ones) defined on $\mathcal{D}$. Note that every associative algebra is automatically a Lie algebra by defining the Lie product proportional to the commutator (cf. Appendix 2):

$$[X, Y] := X \cdot Y - Y \cdot X\,. \tag{21}$$

Since the commutator of two symmetric operators is antisymmetric, we obtain a Lie-algebra structure on the real vector space of symmetric operators with invariant dense domain $\mathcal{D}$ by defining the Lie product as imaginary multiple of the commutator; this I will write as $\frac{1}{\mathrm{i}\hbar}[\cdot,\cdot]$ where $\hbar$ is a real (dimensionful) constant, eventually to be identified with Planck's constant divided by 2π.

Note that I deliberately did *not* state that classical observables should be mapped to *self adjoint* operators. Instead I only required the operators to be symmetric, which is a weaker requirement. This important distinction (see e.g. [14]) is made for the following reason (see e.g. Sect. VIII in [14] for the mathematical distinction): let $\hat{f}$ be the operator corresponding to the phase-space function f. If $\hat{f}$ were self adjoint, then the quantum flow $U(t) = \exp(it\hat{f})$ existed for all $t \in \mathbb{R}$, even if the classical Hamiltonian vector field for f is incomplete (cf. Appendix 1) so that the classical flow does not exist for all flow parameters in $\mathbb{R}$. Hence self adjointness seems too strong a requirement for such f whose classical flow is incomplete (which is the generic situation). Therefore one generally only requires the operators to be symmetric and strengthens this explicitly for those f whose classical flow is complete (see below).

A first attempt to mathematically define Dirac's quantisation strategy could now consist in the following: find a 'suitable' Lie homomorphism $\mathcal{Q}$ from a 'suitable' Lie subalgebra $\mathcal{F}' \subset \mathcal{F}(+,\{,\})$ to the Lie algebra SYM($\mathcal{H}$) of symmetric operators on a Hilbert space $\mathcal{H}$ with some common dense domain $\mathcal{D} \subset \mathcal{H}$. The map $\mathcal{Q}$ will be called the *quantisation map*. Note that this map is a priori not required in any way to preserve the associative structure, i.e. no statement is made to the effect that $\mathcal{Q}(f \cdot g) = \mathcal{Q}(f) \cdot \mathcal{Q}(g)$, or similar.

To be mathematically precise, we still need to interpret the word ‘suitable’ which occurred twice in the above statement. For this we consider the following *test case*, which at first sight appears to be sufficiently general and sufficiently precise to be able to incorporate Dirac’s ideas in a well defined manner:

1. We restrict the Lie algebra of C^∞-Functions on P to polynomials in (q^i, p_i), i.e. we consider $\mathcal{F}_{\text{pol}}(+, \{,\})$.
2. As Hilbert space of states, $\mathcal{H}$, we consider the space of square-integrable functions $\mathbb{R}^n \to \mathfrak{H}$, where $\mathfrak{H}$ is a *finite* dimensional Hilbert space which may account for internal degrees of freedom, like spin. $\mathbb{R}^n$ should be thought of as ‘half’ of phase space, or more precisely the configuration space coordinatised by the set $\{q^1, \cdots, q^n\}$. For integration we take the Lebesgue measure $d^n q$.
3. There exists a map $\mathcal{Q} : \mathcal{F}_{\text{pol}} \to \text{SYM}(\mathcal{H}, \mathcal{D})$ into the set of symmetric operators on $\mathcal{H}$ with common invariant dense domain $\mathcal{D}$. (When convenient we also write $\hat{f}$ instead of $\mathcal{Q}(f)$.) This map has the property that whenever $f \in \mathcal{F}_{\text{pol}}$ has a complete Hamiltonian vector field the operator $Q(f)$ is in fact (essentially) self adjoint.[3]
4. $\mathcal{Q}$ is linear:
$$\mathcal{Q}(f + \lambda\, g) = \mathcal{Q}(f) + \lambda\, \mathcal{Q}(g)\,. \tag{22}$$
5. $\mathcal{Q}$ intertwines the Lie structure on $\mathcal{F}_{\text{pol}}(+, \{,\})$ and the Lie structure given by $\frac{1}{\mathrm{i}\hbar}[,]$ on $\text{SYM}(\mathcal{H}, \mathcal{D})$:
$$\mathcal{Q}(\{f, g\})) = \tfrac{1}{\mathrm{i}\hbar}[\mathcal{Q}(f), \mathcal{Q}(g)])\,. \tag{23}$$
Here $\hbar$ is a constant whose physical dimension is that of $p \cdot q$ (i.e. an action) which accounts for the intrinsic dimension of $\{,\}$ acquired through the differentiations (cf. (6)). Note again that the imaginary unit is necessary to obtain a Lie structure on the subset of symmetric operators.
6. Let 1 also denote the constant function with value 1 on P and $\mathbb{1}$ the unit operator; then
$$\mathcal{Q}(1) = \mathbb{1}\,. \tag{24}$$
7. The quantisation map $\mathcal{Q}$ is consistent with Schrödinger quantisation:
$$\begin{aligned} (\mathcal{Q}(q^i)\psi)(q) &= q^i \psi(q)\,, && (25)\\ (\mathcal{Q}(p_i)\psi)(q) &= -\mathrm{i}\hbar\partial_{q^i}\psi(q)\,. && (26) \end{aligned}$$

[3] We remark that the subset of functions whose flows are complete do not form a Lie subalgebra; hence it would not make sense to just restrict to them.

One might wonder what is actually implied by the last condition and whether it is not unnecessarily restrictive. This is clarified by the theorem of *Stone* and *von Neumann* (see e.g. [1]), which says that if the $2n$ operators $\mathcal{Q}(q^i)$ and $\mathcal{Q}(p_i)$ are represented *irreducibly up to finite multiplicity* (to allow for finitely many internal quantum numbers) and satisfy the required commutation relations, then their representation is unitarily equivalent to the Schrödinger representation given above. In other words, points 2.) and 7.) above are equivalent to, and could therefore be replaced by, the following requirement:

7'. The $2n$ operators $\mathcal{Q}(q^i)$, $\mathcal{Q}(p_i)$ act irreducibly up to at most finite multiplicity on $\mathcal{H}$.

Finally there is a technical point to be taken care of. Note that the commutator on the right hand side of (23) – and hence the whole equation – only makes sense on the subset $\mathcal{D} \subseteq \mathcal{H}$. This becomes important if one deduces from (22) and (23) that

$$\{f, g\} = 0 \Rightarrow [\mathcal{Q}(f), \mathcal{Q}(g)] = 0\,, \tag{27}$$

i.e. that $\mathcal{Q}(f)$ and $\mathcal{Q}(g)$ commute *on* $\mathcal{D}$. Suppose that the Hamiltonian vector fields of f and g are complete so that $Q(f)$ and $Q(g)$ are self adjoint. Then commutativity on $\mathcal{D}$ does *not* imply that $\mathcal{Q}(f)$ and $\mathcal{Q}(g)$ commute in the usual (strong) sense of commutativity of self-adjoint operators, namely that all their spectral projectors mutually commute (compare [14], p. 271). This we pose as an extra condition:

8. If f, g have complete Hamiltonian vector fields and $\{f, g\} = 0$; then $\mathcal{Q}(f)$ commutes with $\mathcal{Q}(g)$ in the strong sense, i.e. their families of spectral projectors commute.

This extra condition will facilitate the technical presentation of the following arguments, but we remark that it can be dispensed with [8].

2.3 The Theorem of Groenewold and van Howe

In a series of papers Groenewold [9] and van Hove [16,15] showed that a canonical quantisation satisfying requirements 1.–8. does *not* exist. The proof is instructive and therefore we shall present it in detail. For logical clarity it is advantageous to divide it into two parts:

Part 1 shows the following 'squaring laws':

$$\begin{aligned} \mathcal{Q}(q^2) &= [\mathcal{Q}(q)]^2\,, & (28)\\ \mathcal{Q}(p^2) &= [\mathcal{Q}(p)]^2\,, & (29)\\ \mathcal{Q}(qp) &= \tfrac{1}{2}[\mathcal{Q}(q)\mathcal{Q}(p) + \mathcal{Q}(p)\mathcal{Q}(q)]\,. & (30) \end{aligned}$$

Next to elementary manipulations the proof of part 1 uses a result concerning the Lie algebra $sl(2, \mathbb{R})$, which we shall prove in Appendix 2. Note that in the canonical approach as formulated here *no* initial assumption whatsoever was made concerning the preservation of the associative structure. Points 4. and 5. only required the Lie structure to the preserved. The importance of part 1 is to show that such a partial preservation of the associative structure can actually be derived. It will appear later (cf. Sect. 2.5) that this consequence could not have been drawn without the irreducibility requirement 7').

Part 2 shows that the squaring laws lead to a contradiction to (23) on the level of higher than second-order polynomials.

Let us now turn to the proofs. To save notation we write $\hat{f}$ instead of $\mathcal{Q}(f)$. Also, we restrict attention to $n = 1$, i.e. we have one q and one p coordinate on the two dimensional phase space $\mathbb{R}^2$. In what follows, essential use is repeatedly made of condition 8 in the following form: assume $\{f, q\} = 0$ then (23) and condition 8 require that $\hat{f}$ (strongly) commutes with $\hat{q}$, which in the Schrödinger representation implies that $\hat{f}$ has the form $(\hat{f}\psi)(q) = A(q)\psi(q)$, where $A(q)$ is a Hermitean operator (matrix) in the finite dimensional (internal) Hilbert space $\mathfrak{H}$.

Proof of Part 1. We shall present the argument in 7 small steps. Note that throughout we work in the Schrödinger representation.

i) Calculate $\widehat{q^2}$: we have $\{q^2, q\} = 0$, hence $\widehat{q^2} = A(q)$. Applying (23) and (25) to $\{p, q^2\} = -2q$ gives $\frac{1}{i\hbar}[\hat{p}, \widehat{q^2}] = -2\hat{q}$ and hence $A'(q) = 2q$ (here we suppress to write an explicit $\mathbb{1}$ for the unit operator in $\mathfrak{H}$), so that

$$\widehat{q^2} = \hat{q}^2 - 2\mathfrak{e}_-\,, \tag{31}$$

where $\mathfrak{e}_-$ is a constant (i.e. q independent) Hermitean matrix in $\mathfrak{H}$.

ii) Calculate $\widehat{p^2}$: this is easily obtained by just Fourier transforming the case just done. Hence

$$\widehat{p^2} = \hat{p}^2 + 2\mathfrak{e}_+\,, \tag{32}$$

where $\mathfrak{e}_+$ is a constant Hermitean matrix in $\mathfrak{H}$ (here, as in (31), the conventional factor of 2 and the signs are chosen for later convenience).

iii) Calculate $\widehat{qp}$: We apply (23) to $4qp = \{q^2, p^2\}$ and insert the results (31) and (32):

$$\widehat{qp} = \tfrac{1}{4i\hbar}[\widehat{q^2}, \widehat{p^2}] = \tfrac{1}{4i\hbar}[\hat{q}^2, \hat{p}^2] - \tfrac{1}{i\hbar}[\mathfrak{e}_-, \mathfrak{e}_+] = \tfrac{1}{2}(\hat{q}\hat{p} + \hat{p}\hat{q}) + \mathfrak{h}\,, \tag{33}$$

where

$$\mathfrak{h} := \tfrac{1}{i\hbar}[\mathfrak{e}_+, \mathfrak{e}_-]\,. \tag{34}$$

In the last step of (33) we iteratively used the general rule

$$[A, BC] = [A, B]C + B[A, C]\,. \tag{35}$$

iv) Next consider the quantities

$$
\begin{aligned}
h &:= \tfrac{1}{2}(\hat{q}\hat{p}+\hat{p}\hat{q})\,, && (36)\\
e_+ &:= \tfrac{1}{2}\hat{p}^2\,, && (37)\\
e_- &:= -\tfrac{1}{2}\hat{q}^2\,. && (38)
\end{aligned}
$$

By straightforward iterative applications of (35) short computations yield

$$\tfrac{1}{\mathrm{i}\hbar}[e_+,e_-]=h\,,\quad \tfrac{1}{\mathrm{i}\hbar}[h,e_\pm]=\pm 2\,e_\pm\,, \tag{39}$$

which show that $e_\pm, h$ furnish a representation of the Lie algebra of $sl(2,\mathbb{R})$ of real traceless 2×2 matrices (see Appendix 2 for details).

v) On the other hand, defining

$$
\begin{aligned}
H &:= \widehat{qp}\,, && (40)\\
E_+ &:= \tfrac{1}{2}\widehat{p^2}\,, && (41)\\
E_- &:= -\tfrac{1}{2}\widehat{q^2}\,, && (42)
\end{aligned}
$$

we can directly use (23) to calculate their Lie brackets. This shows that they also satisfy the $sl(2,\mathbb{R})$ algebra:

$$\tfrac{1}{\mathrm{i}\hbar}[E_+,E_-]=H\,,\quad \tfrac{1}{\mathrm{i}\hbar}[H,E_\pm]=\pm 2\,E_\pm\,. \tag{43}$$

vi) Inserting into (43) the results (31-33) now implies that the Hermitean matrices $\mathfrak{e}_\pm, \mathfrak{h}$ too satisfy the $sl(2,\mathbb{R})$ algebra:

$$\tfrac{1}{\mathrm{i}\hbar}[\mathfrak{e}_+,\mathfrak{e}_-]=\mathfrak{h}\,,\quad \tfrac{1}{\mathrm{i}\hbar}[\mathfrak{h},\mathfrak{e}_\pm]=\pm 2\,\mathfrak{e}_\pm\,. \tag{44}$$

vii) Finally we invoke the following result from Appendix 2:

LEMMA 1. *Let A, B_+, B_- be finite dimensional anti-Hermitean matrices which satisfy $A=[B_+,B_-]$ and $[A,B_\pm]=\pm 2B_\pm$, then $A=B_\pm=0$.*

Applying this to our case by setting $A=\frac{1}{\mathrm{i}\hbar}\mathfrak{h}$ and $B_\pm=\frac{1}{\mathrm{i}\hbar}\mathfrak{e}_\pm$ implies

$$\mathfrak{e}_\pm = 0 = \mathfrak{h}\,. \tag{45}$$

Inserting this into (31-33) yields (28-30) respectively. This ends the proof of part 1.

Proof of Part 2. Following [8], we first observe that the statements (28-30) can actually be generalised: Let P be any real polynomial, then

$$
\begin{aligned}
\widehat{P(q)} &= P(\hat{q})\,, && (46)\\
\widehat{P(p)} &= P(\hat{p})\,, && (47)\\
\widehat{P(q)p} &= \tfrac{1}{2}(P(\hat{q})\hat{p}+\hat{p}P(\hat{q}))\,, && (48)\\
\widehat{P(p)q} &= \tfrac{1}{2}(P(\hat{p})\hat{q}+\hat{q}P(\hat{p}))\,. && (49)
\end{aligned}
$$

To complete the proof of part 2 it is sufficient to prove (46) and (47) for $P(x) = x^3$, and (48) and (49) for $P(x) = x^2$. This we shall do first. The cases for general polynomials – which we do not need – follow by induction and linearity. Again we break up the argument, this time into 5 pieces.

i) We first note that $\{q, q^3\} = 0$ implies via (23) that $\hat{q}$ and $\widehat{q^3}$ commute. Since $\hat{q}$ and $\hat{q}^3$ commute anyway we can write $\widehat{q^3} - \hat{q}^3 = A(q)$, where $A(q)$ takes values in the space of Hermitean operators on $\mathfrak{H}$.
ii) We next show that $A(q)$ also commutes with $\hat{p}$. This follows from the following string of equations, where we indicated the numbers of the equations used in the individual steps as superscripts over the equality signs:
$$[\widehat{q^3}, \hat{p}] \stackrel{23}{=} \mathrm{i}\hbar\widehat{\{q^3, p\}} \stackrel{6}{=} 3\mathrm{i}\hbar\widehat{q^2} \stackrel{28}{=} 3\mathrm{i}\hbar\hat{q}^2 \stackrel{35}{=} [\hat{q}^3, \hat{p}]\,. \tag{50}$$
Hence $A(q)$ equals a q-independent matrix, $\mathfrak{a}$, and we have
$$\widehat{q^3} = \hat{q}^3 + \mathfrak{a}\,. \tag{51}$$
iii) We show that the matrix $\mathfrak{a}$ must actually be zero by the following string of equations:
$$\begin{aligned}\widehat{q^3} \stackrel{6}{=} \tfrac{1}{3}\widehat{\{q^3, qp\}} \stackrel{23}{=} \tfrac{1}{3\mathrm{i}\hbar}[\widehat{q^3}, \widehat{qp}] &\stackrel{30,51}{=} \tfrac{1}{3\mathrm{i}\hbar}[\hat{q}^3 + \mathfrak{a}, \tfrac{1}{2}(\hat{q}\hat{p} + \hat{p}\hat{q})] \\ &\stackrel{*}{=} \tfrac{1}{6\mathrm{i}\hbar}[\hat{q}^3, (\hat{q}\hat{p} + \hat{p}\hat{q})] \stackrel{35}{=} \hat{q}^3\,, \end{aligned} \tag{52}$$
where at $*$ we used that $\mathfrak{a}$ commutes with $\hat{q}$ and $\hat{p}$. This proves (46) for $P(q) = q^3$. Exchanging p and q and repeating the proof shows (47) for $P(p) = p^3$.
iv) Using what has been just shown allows to prove (48) for $P(q) = q^2$:
$$\widehat{q^2p} \stackrel{6}{=} \tfrac{1}{6}\widehat{\{q^3, p^2\}} \stackrel{23}{=} \tfrac{1}{6\mathrm{i}\hbar}[\widehat{q^3}, \widehat{p^2}] \stackrel{46,29}{=} \tfrac{1}{6\mathrm{i}\hbar}[\hat{q}^3, \hat{p}^2] \stackrel{35}{=} \tfrac{1}{2}(\hat{q}^2\hat{p} + \hat{p}\hat{q}^2)\,. \tag{53}$$
Exchanging q and p proves (49) for $P(p) = p^2$.
v) Finally we apply the quantisation map to both sides of the classical equality
$$\tfrac{1}{9}\{q^3, p^3\} = \tfrac{1}{3}\{q^2p, p^2q\}\,. \tag{54}$$
On the left hand side we replace $\widehat{q^3}$ and $\widehat{p^3}$ with $\hat{q}^3$ and $\hat{q}^3$ respectively and then successively apply (35); this leads to
$$\hat{q}^2\hat{p}^2 - 2\mathrm{i}\hbar\hat{q}\hat{p} - \tfrac{2}{3}\hbar^2\mathbb{1}\,. \tag{55}$$
On the right hand side of (54) we now use (48) and (49) to replace $\widehat{q^2p}$ and $\widehat{p^2q}$ with $\frac{1}{2}(\hat{q}^2\hat{p}+\hat{p}\hat{q}^2)$ and $\frac{1}{2}(\hat{p}^2\hat{q}+\hat{q}\hat{p}^2)$ respectively and again successively apply (35). This time we obtain
$$\hat{q}^2\hat{p}^2 - 2\mathrm{i}\hbar\hat{q}\hat{p} - \tfrac{1}{3}\hbar^2\mathbb{1}\,, \tag{56}$$
which differs from (55) by a term $-\frac{1}{3}\hbar^2\mathbb{1}$. But according to (23) both expressions should coincide, which means that we arrived at a contradiction. This completes part 2 and hence the proof of the theorem of Groenewold and van Howe.

2.4 Discussion

The GvH-Theorem shows that the Lie algebra of *all* polynomials on $\mathbb{R}^{2n}$ cannot be quantised (and hence no Lie subalgebra of $C^\infty(P)$ containing the polynomials). But its proof has also shown that the Lie *sub*algebra

$$\mathcal{F}_{\text{pol}(2)} := \text{span}\left\{1, q, p, q^2, p^2, qp\right\} \tag{57}$$

of polynomials of at most quadratic order *can* be quantised. This is just the essence of the 'squaring laws' (28-30).

To see that $\mathcal{F}_{\text{pol}(2)}$ is indeed a Lie subalgebra, it is sufficient to note that the Poisson bracket (6) of a polynomial of n-th and a polynomial of m-th order is a polynomial of order $(n+m-2)$. Moreover, it can be shown that $\mathcal{F}_{\text{pol}(2)}$ is a *maximal* Lie subalgebra of $\mathcal{F}_{\text{pol}}$, i.e. that there is no other proper Lie subalgebra $\mathcal{F}'$ which properly contains $\mathcal{F}_{\text{pol}(2)}$, i.e. which satisfies $\mathcal{F}_{\text{pol}(2)} \subset \mathcal{F}' \subset \mathcal{F}_{\text{pol}}$.

$\mathcal{F}_{\text{pol}(2)}$ contains the Lie subalgebra of all polynomials of at most first order:

$$\mathcal{F}_{\text{pol}(1)} := \text{span}\{1, q, p\} \; . \tag{58}$$

This is clearly a Lie ideal in $\mathcal{F}_{\text{pol}(2)}$ (not in $\mathcal{F}_{\text{pol}}$), since Poisson brackets between quadratic and linear polynomials are linear. $\mathcal{F}_{\text{pol}(1)}$ is also called the 'Heisenberg algebra'. According to the rules (25,26) the Heisenberg algebra was required to be represented irreducibly (cf. the discussion following (26)). What is so special about the Heisenberg algebra? First, observe that it contains enough functions to coordinatise phase space, i.e. that no two points in phase space assign the same values to the functions contained in the Heisenberg algebra. Moreover, it is a minimal subalgebra of $\mathcal{F}_{\text{pol}}$ with this property. Hence it is a minimal set of classical observables whose values allow to uniquely fix a classical state (point in phase space). The irreducibility requirement can then be understood as saying that this property should essentially also be shared by the quantised observables, at least up to *finite* multiplicities which correspond to the 'internal' Hilbert space $\mathfrak{H}$ (a ray of which is fixed by finitely many eigenvalues). We will have more to say about this irreducibility postulate below.

The primary lesson from the GvH is that $\mathcal{F}_{\text{pol}} \subset \mathcal{F}_\infty$ was chosen too big. It is not possible to find a quantisation map $\mathcal{Q} : \mathcal{F}_{\text{pol}}(+, \{,\}) \to \text{SYM}(\mathcal{H})$ which intertwines the Lie structures $\{,\}$ and $\frac{1}{i\hbar}[,]$. This forces us to reformulate the canonical quantisation programme. From the discussion so far one might attempt the following rules

Rule 1. Given the Poisson algebra $\mathcal{F}_{\text{pol}}(+, \{,\}, \cdot)$ of all polynomials on phase space. Find a Lie subalgebra $\mathcal{F}_{\text{irr}} \subset \mathcal{F}_{\text{pol}}(+, \{,\})$ of 'basic observables' which fulfills the two conditions: (1) $\mathcal{F}_{\text{irr}}$ contains sufficiently many functions so as to coordinatise phase space, i.e. no two points coincide in all values of functions in $\mathcal{F}_{\text{irr}}$; (2) $\mathcal{F}_{\text{irr}}$ is minimal in that respect, i.e. there is no Lie subalgebra $\mathcal{F}'_{\text{irr}}$ properly contained in $\mathcal{F}_{\text{irr}}$ which also fulfills (1).

Rule 2. Find another Lie subalgebra $\mathcal{F}_{\text{quant}} \subset \mathcal{F}_{\text{pol}}(+, \{,\})$ so that $\mathcal{F}_{\text{irr}} \subseteq \mathcal{F}_{\text{quant}}$ and that $\mathcal{F}_{\text{quant}}$ can be quantised, i.e. a Lie homomorphism $\mathcal{Q} : \mathcal{F}_{\text{quant}} \to \text{SYM}(\mathcal{H})$

can be found, which intertwines the Lie structures $\{,\}$ and $\frac{1}{i\hbar}[,]$. Require $\mathcal{Q}$ to be such that $\mathcal{Q}(\mathcal{F}_{\rm irr})$ act almost irreducibly, i.e. up to finite multiplicity, on $\mathcal{H}$. Finally, require that $\mathcal{F}_{\rm quant}$ be maximal in $\mathcal{F}_{\rm pol}$, i.e. that there is no $\mathcal{F}'_{\rm quant}$ with $\mathcal{F}_{\rm quant} \subset \mathcal{F}'_{\rm quant} \subset \mathcal{F}_{\rm pol}(+,\{,\})$.

Note that the choice of $\mathcal{F}_{\rm quant}$ is generally far from unique. For example, instead of choosing $\mathcal{F}_{\rm quant} = \mathcal{F}_{\rm pol(2)}$, i.e. the polynomials of at most quadratic order, we could choose $\mathcal{F}_{\rm quant} = \mathcal{F}_{\rm pol(\infty,1)}$, the polynomials of at most linear order in momenta with coefficients which are arbitrary polynomials in q. A general element in $\mathcal{F}_{\rm pol(\infty,1)}$ has the form

$$f(q,p) = g(q) + h(q)\,p \tag{59}$$

where g, h are arbitrary polynomials with real coefficients. The Poisson bracket of two such functions is

$$\{f_1, f_2\} = \{g_1 + h_1 p, g_2 + h_2 p\} = g_3 + h_3 p\,, \tag{60}$$

where

$$g_3 = g_1' h_2 - g_2' h_1 \quad \text{and} \quad h_3 = h_1' h_2 - h_1 h_2'\,. \tag{61}$$

The quantisation map applied to f is then given by

$$\widehat{f} = g(\hat{q}) - i\hbar\big(\tfrac{1}{2}h'(\hat{q}) + h(\hat{q})\,\tfrac{d}{dq}\big)\,, \tag{62}$$

where h' denotes the derivative of h and $\hat{q}$ and $\hat{p}$ are just the Schrödinger operators 'multiplication by q' and '$-i\hbar d/dq$' respectively. The derivative term proportional to h' is necessary to make $\widehat{f}$ symmetric (an overline denoting complex conjugation):

$$\begin{aligned}\big[\tfrac{i}{2}h'(q)\psi(q) + ih(q)\psi'(q)\big]\,\overline{\phi(q)} \;&=\; \psi(q)\,\overline{\big[\tfrac{i}{2}h'(q)\phi(q) + ih(q)\phi'(q)\big]} \\ &\quad + \;(ih\psi\overline{\phi})'(q)\,,\end{aligned} \tag{63}$$

where the last term vanishes upon integration. Moreover, a simple computation readily shows that the map $f \mapsto \hat{f}$ indeed defines a Lie homomorphism from $\mathcal{F}_{\rm pol(\infty,1)}$ to $\mathrm{SYM}(\mathcal{H})$:

$$\tfrac{1}{i\hbar}[\hat{f}_1, \hat{f}_2] = g_3(q) - i\hbar\left(\tfrac{1}{2}h_3'(q) + h_3(q)\tfrac{d}{dq}\right) = \widehat{\{f_1, f_2\}}\,, \tag{64}$$

with $f_{1,2}$ and g_3, h_3 as in (60) and (61) respectively. Hence (62) gives a quantisation of $\mathcal{F}_{\rm pol(\infty,1)}$.

It can be shown ([8], Theorem 8) that $\mathcal{F}_{\rm pol(2)}$ and $\mathcal{F}_{\rm pol(\infty,1)}$ are the only maximal Lie subalgebras of $\mathcal{F}_{\rm pol}$ which contain the Heisenberg algebra $\mathcal{F}_{\rm pol(1)}$. In this sense, if one restricts to polynomial functions, there are precisely two inextendible quantisations.

So far we restricted attention to polynomial functions. Since $\mathcal{F}_{\text{pol}}$ is already too big to be quantised, there is clearly no hope to quantise all C^∞ functions on our phase space $\mathbb{R}^{2n}$. For general phase spaces P (i.e. not isomorphic to $\mathbb{R}^{2n}$) there is generally no notion of 'polynomials' and hence no simple way to characterise suitable Lie subalgebras of $\mathcal{F}_\infty(+, \{,\})$. But experience with the GvH Theorem suggests anyway to conjecture that, subject to some irreducibility postulate for some minimal choice of $\mathcal{F}_{\text{irr}} \subset \mathcal{F}_\infty$, there is *never* a quantisation of $\mathcal{F}_\infty$. (A quantisation of all C^∞ functions is called *full quantisation* in the literature.) Surprisingly there is a non-trivial counterexample to this conjecture: it has been shown that a full quantisation exists for the 2-torus [6]. One might first guess that this is somehow due to the compactness of the phase space. But this is not true, as a GvH obstruction to full quantisation does exist for the 2-sphere [7]. But the case of the 2-torus seems exceptional, even mathematically. The general expectation is indeed that GvH-like obstructions are in some sense generic, though, to my knowledge, there is no generally valid formulation and corresponding theorem to that effect. (For an interesting early attempt in this direction see [5].) Hence we face the problem to determine $\mathcal{F}_{\text{irr}}$ and $\mathcal{F}_{\text{quant}}$ within $\mathcal{F}_\infty$. There is no general theory how to do this. If P is homogeneous, i.e. if there is a finite dimensional Lie group G (called the 'canonical group') that acts transitively on P and preserves the Poisson bracket (like the $2n$ translations in $\mathbb{R}^{2n}$) one may generate $\mathcal{F}_{\text{irr}}$ from the corresponding momentum maps. This leads to a beautiful theory [12] for such homogeneous situations, but general finite dimensional P do not admit a finite dimensional canonical group G, and then things become much more complicated.

2.5 The Role of the Irreducibility-Postulate

In this section we wish to point out the central role played by the irreducibility postulate. We already mentioned in Sect. 2.3 that the GvH theorem could not have been derived without it. Let us show this by dropping that postulate and see what happens. This leads to a weaker notion of quantisation which deserves to be considered in its own right:

Definition 2. Quantisation without the irreducibility postulate (25,26) is called *pre-quantisation.*

Given the GvH result, the following is remarkable:

Theorem 1. *A prequantisation of the Lie algebra $\mathcal{F}_\infty(+, \{,\})$ of all C^∞-functions on $\mathbb{R}^{2n}$ exists.*

The proof is constructive by means of *geometric quantisation.* Let us briefly recall the essentials of this approach: The Hilbert space of states is taken to be $\mathcal{H} = L^2(\mathbb{R}^{2n}, d^nq d^np)$, i.e. the square integrable functions on *phase* space ($2n$ coordinates), instead of configuration space (n coordinates). The quantisation map is as follows[4]:

$$\mathcal{Q}(f) = \mathrm{i}\hbar\nabla_{X_f} + f\,, \tag{66}$$

where ∇ is a 'covariant-derivative' operator, which is

$$\nabla = d + A\,. \tag{67}$$

Here d is just the ordinary (exterior) derivative and the connection 1-form, A, is proportional to the canonical 1-form (cf. (96)) $\theta := p_i\, dq^i$:

$$A = -\tfrac{\mathrm{i}}{\hbar}\,\theta = -\tfrac{\mathrm{i}}{\hbar}p_i\, dq^i\,. \tag{68}$$

The curvature, $F = dA$, is then proportional to the symplectic 2-form $\omega = d\theta$:

$$F = \tfrac{\mathrm{i}}{\hbar}\omega = \tfrac{\mathrm{i}}{\hbar}dq^i \wedge dp_i\,. \tag{69}$$

If X_f is the Hamiltonian vector field on phase space associated to the phase-space function f (cf. (91)), then in canonical coordinates it has the form

$$X_f = (\partial_{p_i} f)\partial_{q^i} - (\partial_{q^i} f)\partial_{p_i}\,. \tag{70}$$

The map $f \mapsto X_f$ is a Lie homomorphism from $\mathcal{F}_\infty(+,\{,\})$ to the Lie algebra of vector fields on phase space, i.e. $X_{\{f,g\}} = [X_f, X_g]$. The operator $\hat{f}$ is formally self-adjoint and well defined on Schwarz-space (rapidly decreasing functions), which we take as our invariant dense domain $\mathcal{D}$. Explicitly its action reads:

$$\hat{f} = i\hbar\big((\partial_{q^i} f)\partial_{p_i} - (\partial_{p_i} f)\partial_{q^i}\big) + \big(f - (\partial_{p_i} f)p_i\big)\,, \tag{71}$$

which clearly shows that all operators are differential operators of at most degree one. This makes it obvious that a squaring-law in the form $\hat{f}\hat{g} = \widehat{fg}$ *never* applies. For example, for $n = 1$ we have for $\hat{q}, \hat{p}$ and their squares:

$$\hat{q} = q + \mathrm{i}\hbar\partial_p\,, \quad \widehat{q^2} = q^2 + 2\mathrm{i}\hbar\,\partial_p\,, \tag{72}$$

$$\hat{p} = -\mathrm{i}\hbar\partial_q\,, \quad \widehat{p^2} = -p^2 - 2\mathrm{i}\hbar\, p\partial_q\,. \tag{73}$$

[4] Unlike in ordinary Schrödinger quantisation, where $|\psi(q)|^2$ is the probability density for the system in configuration space, the corresponding quantity $|\psi(q,p)|^2$ in geometric quantisation has *not* the interpretation of a probability density in phase space. The formal reason being that in geometric quantisation $\hat{q}$ is not just a multiplication operator (cf. (72)). For example, if ψ has support in an arbitrary small neighbourhood U of phase space this does not mean that we can simultaneously reduce the uncertainties of $\hat{q}$ and $\hat{p}$, since this would violate the uncertainty relations which hold unaltered in geometric quantisation. Recall that the uncertainty relations just depend on the commutation relations since they derive from the following generally valid formula by dropping the last term: ($\langle\cdot\rangle_\psi$ denotes the expectation value in the state ψ, $[\cdot,\cdot]_+$ the anticommutator and $\hat{f}_0 := \hat{f} - \langle\hat{f}\rangle_\psi \mathbb{1}$):

$$\langle\hat{f}_0^2\rangle_\psi\langle\hat{g}_0^2\rangle_\psi \geq \tfrac{1}{4}\left\{|\langle[\hat{f},\hat{g}]\rangle_\psi|^2 + |\langle[\hat{f}_0,\hat{g}_0]_+\rangle_\psi|^2\right\}\,. \tag{65}$$

One now proves by direct computation that (66) indeed defines a Lie homomorphism:

$$\begin{aligned}\frac{1}{\mathrm{i}\hbar}\left[\mathcal{Q}(f),\mathcal{Q}(g)\right] &= \frac{1}{\mathrm{i}\hbar}\left[\mathrm{i}\hbar\nabla_{X_f}+f\,,\,\mathrm{i}\hbar\nabla_{X_g}+g\right]\\ &= \mathrm{i}\hbar\left[\nabla_{X_f},\nabla_{X_g}\right]+X_f(g)-X_g(f)\\ &= \mathrm{i}\hbar\left(\nabla_{[X_f,X_g]}+F(X_f,X_g)\right)+2\{f,g\}\\ &= \mathrm{i}\hbar\,\nabla_{X_{\{f,g\}}}+\{f,g\}=\mathcal{Q}(\{f,g\})\,,\end{aligned} \tag{74}$$

where we just applied the standard identity for the curvature of the covariant derivative (67): $F(X,Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}$ and also used $-\mathrm{i}\hbar\,F(X_f,X_g)=\omega(X_f,X_g)=\{f,g\}$ (cf. (94)).

Let us now look at a simple specific example: the linear harmonic oscillator. We use units where its mass and angular frequency equal 1. The Hamiltonian function and vector field are then given by:

$$H=\tfrac{1}{2}(p^2+q^2)\;\Rightarrow\;X_H=p\partial_q-q\partial_p\,, \tag{75}$$

whose quantisation according to (66) is

$$\hat{H}=-\mathrm{i}\hbar\,(p\partial_q-q\partial_p)+\tfrac{1}{2}\left(q^2-p^2\right)\,. \tag{76}$$

Introducing polar coordinates on phase space: $q = r\cos(\varphi)$ $p = r\sin(\varphi)$, the Hamiltonian becomes

$$\hat{H}=\mathrm{i}\hbar\partial_\varphi+\tfrac{r^2}{2}\cos(2\varphi)\,. \tag{77}$$

The eigenvalue equation reads

$$\hat{H}\psi=E\psi\;\Leftrightarrow\;\partial_\varphi\psi=-\tfrac{\mathrm{i}}{\hbar}\left(E-\tfrac{r^2}{2}\cos(2\varphi)\right)\,\psi\,, \tag{78}$$

whose solution is

$$\psi(r,\varphi)=\psi_0(r)\,\exp\left\{-\tfrac{\mathrm{i}}{\hbar}\left(E\varphi-\tfrac{r^2}{2}\,\sin(2\varphi)\right)\right\}\,, \tag{79}$$

where ψ_0 is an arbitrary function in $L^2(\mathbb{R}_+,rdr)$. Single valuedness requires

$$E=E_n=n\hbar\,,\quad n\in\mathbb{Z}\,, \tag{80}$$

with each energy eigenspace being isomorphic to the space of square-integrable functions on the positive real line with respect to the measure rdr:

$$\mathcal{H}_n=L^2(\mathbb{R}_+,rdr)\,. \tag{81}$$

Hence we see that the difference to the usual Schrödinger quantisation is not simply an expected degeneracy of the energy eigenspaces which, by the way, turns out to be quite enormous, i.e. infinite dimensional for each energy level. What is much worse and perhaps less expected is the fact that the energy spectrum in prequantisation is a proper extension of that given by Schrödinger quantisation and, in distinction to the latter, that it is *unbounded from below*. This means that there is no ground state for the harmonic oscillator in prequantisation which definitely appears physically wrong. Hence there seems to be some deeper physical significance to the irreducibility postulate than just mere avoidance of degeneracies.

3 Constrained Systems

For systems with gauge redundancies[5] the original phase space P does not directly correspond to the set of (mutually different) classical states. First of all, only a subset $\hat{P} \subset P$ will correspond to classical states of the system, i.e. the system is *constrained* to $\hat{P}$. Secondly, the points of $\hat{P}$ label the states of the systems in a redundant fashion, that is, one state of the classical system is labeled by many points in $\hat{P}$. The set of points which label the same state form an orbit of the group of gauge transformations which acts on $\hat{P}$. 'Lying in the same orbit' defines an equivalence relation (denoted by $\sim$) on $\hat{P}$ whose equivalence classes form the space $\bar{P} := \hat{P}/\sim$ which is called the *reduced phase space*. Its points now label the classical states in a faithful fashion. Note that it is a quotient-space of the sub-space $\hat{P}$ of P and can, in general, therefore not be represented as a subspace of P.

A straightforward strategy to quantise such a system is to 'solve' the constraints, that is, to construct $\bar{P}$. One could then apply the same methods as for unconstrained systems, at least as long as $\bar{P}$ will be a C^∞-manifold with a symplectic structure (cf. Appendix 1).[6] In particular, we can then consider the Poisson algebra of C^∞-functions and proceed as for unconstrained systems.

However, in general it is analytically very difficult to explicitly do the quotient construction $\hat{P} \to \hat{P}/\sim\, = \bar{P}$, i.e. to solve the constraints *classically*. Dirac has outlined a strategy to implement the constraints *after* quantisation [4]. The basic mathematical reason why this is considered a simplification is seen in the fact that the whole problem is now posed in *linear* spaces, i.e. the construction of sub- and quotient spaces in the (linear) spaces of states and observables.

Dirac's ideas have been reviewed, refined, and discussed many times in the literature; see e.g. the comprehensive textbook by Henneaux and Teitelboim [11]. Here we shall merely give a brief coordinate-free description of how to construct the right classical Poisson algebra of functions (the 'physical observables').

3.1 First-Class Constraints

Let (P,ω) be a symplectic manifold which is to be thought of as an initial phase space of some gauge system. The physical states then correspond to the points of some submanifold $\hat{P} \hookrightarrow P$. Usually $\hat{P}$ is characterised as zero-level set of some given collection of functions, $\hat{P} = \{p \in P \mid \phi_\alpha(p) = 0,\ \alpha = 1, 2, ..., \mathrm{codim}(\hat{P})\}$, where $\mathrm{codim}(\hat{P}) := \dim(P) - \dim(\hat{P})$ denotes the codimension of $\hat{P}$ in P. The ensuing formulae will then depend on the choice of ϕ_α, though the resulting

[5] We deliberately avoid the word 'symmetry' in this context, since the action of a gauge group has a completely different physical interpretation than the action of a proper symmetry; only the latter transforms states into other, physically distinguishable states. See Sect. 6.3 in [13] for a more comprehensive discussion of this point.

[6] In passing we remark that even though P may (and generally is in applications) a cotangent bundle T^*Q for some configuration space Q, this need not be true for $\bar{P}$, i.e. there will be no space $\bar{Q}$ such that $\bar{P} \cong T^*\bar{Q}$. For this reason it is important to develop quantisations strategies that apply to general symplectic manifolds.

theory should only depend on $\hat{P}$ and not on its analytical characterisation. To make this point manifest we just work with the geometric data. As usual, we shall denote the tangent bundles of P and $\hat{P}$ by TP and $T\hat{P}$ respectively. The restriction of TP to $\hat{P}$ (which also contains vectors not tangent to $\hat{P}$) is given by $TP|_{\hat{P}} := \{X \in T_pP \mid p \in \hat{P}\}$. The ω-orthogonal complement of $T_p\hat{P}$ is now defined as follows:

$$T_p^{\perp}\hat{P} := \{X \in T_pP|_{\hat{P}} \mid \omega(X,Y) = 0,\ \forall Y \in T_pP\}\,. \tag{82}$$

Definition 3. A submanifold $\hat{P} \hookrightarrow P$ is called *coisotropic* iff $T^{\perp}\hat{P} \subset T\hat{P}$.

Since ω is non degenerate we have $\dim T_p\hat{P} + \dim T_p^{\perp}\hat{P} = \dim T_pP$, hence $\dim T_p^{\perp}\hat{P} = \operatorname{codim}\hat{P}$. This means that for coisotropic embeddings $i : \hat{P} \hookrightarrow P$ the kernel[7] of the pulled-back symplectic form $\hat{\omega} := i^*\omega$ on $\hat{P}$ has the maximal possible number of dimensions, namely $\operatorname{codim}\hat{P}$.

Definition 4. A constrained system $\hat{P} \hookrightarrow P$ is said to be of *first class* iff $\hat{P}$ is a coisotropic submanifold of (P,ω).

From now on we consider only first class constraints.

Lemma 2. *$T^{\perp}\hat{P} \subset TP|_{\hat{P}}$ is an integrable subbundle.*

Proof. The statement is equivalent to saying that the commutator of any two $T^{\perp}\hat{P}$-valued vector fields X,Y on $\hat{P}$ is again $T^{\perp}\hat{P}$-valued. Using $[X,Y] = L_XY$ and formula (93) we have[8] $[X,Y] \vdash \hat{\omega} = L_X(Y \vdash \hat{\omega}) - Y \vdash L_X\hat{\omega} = -Y \vdash d(X \vdash \hat{\omega}) = 0$, since $Y \vdash \hat{\omega} = 0 = X \vdash \hat{\omega}$ and $d\hat{\omega} = di^*\omega = i^*d\omega = 0$ due to $d\omega = 0$. □

Definition 5. The *gauge algebra*, Gau, is defined to be the set of all functions (out of some function class $\mathcal{F}$, usually $C^\infty(P)$) which vanish on $\hat{P}$:

$$\mathrm{Gau} := \{f \in \mathcal{F}(P) \mid f|_{\hat{P}} \equiv 0\}\,. \tag{83}$$

Gau uniquely characterises the constraint surface $\hat{P}$ in a coordinate independent fashion. In turn, this allows to characterise the constraints algebraically; Gau is in fact a Poisson algebra. To see this, first note that it is obviously an ideal of the associative algebra $\mathcal{F}(+,\cdot)$, since any pointwise product with an element in Gau also vanishes on $\hat{P}$. Next we show

Lemma 3. $f \in \mathrm{Gau}$ *implies that $X_f|_{\hat{P}}$ is $T^{\perp}\hat{P}$-valued.*

---

[7] The kernel (or 'null-space') of a bilinear form f on V is the subspace $\mathrm{kernel}(f) := \{X \in V \mid f(X,Y) = 0,\ \forall Y \in V\}$.

[8] We shall use the symbol $\vdash$ to denote the insertion of a vector (standing to the left of $\vdash$) into the first slot of a form (standing to the right of $\vdash$). For example, for the 2-form ω, $X \vdash \omega$ denotes the 1-form $\omega(X,\cdot)$.

Proof. $f|_{\hat{P}} \equiv 0 \Rightarrow \text{kernel}(df|_{\hat{P}}) = \text{kernel}((X_f \vdash \omega)|_{\hat{P}}) \supseteq T\hat{P}$. Hence $X_f|_{\hat{P}}$ is $T^{\perp}\hat{P}$-valued. □

Now it is easy to see that Gau is also a Lie algebra, since for $f, g \in \text{Gau}$ we have

$$\{f, g\}|_{\hat{P}} = X_f(g)|_{\hat{P}} = X_f \vdash dg|_{\hat{P}} = X_g \vdash X_f \vdash \omega|_{\hat{P}} = 0\,, \tag{84}$$

where (91) and Lemma 3 was used in the last step. Hence Gau is shown to be an associative and Lie algebra, hence a Poisson algebra. But note that whereas it is an associative ideal it is not a Lie ideal. Indeed, for $f \in \text{Gau}$ and $g \in \mathcal{F}$ we have $\{f, g\}|_{\hat{P}} = X_f(g)|_{\hat{P}} \neq 0$ for those g which vary on $\hat{P}$ in the direction of X_f.

The interpretation of Gau is that its Hamiltonian vector fields generate gauge transformations, that is, motions which do not correspond to physically existing degrees of freedom. Two points in $\hat{P}$ which are on the same connected leaf of $T^{\perp}\hat{P}$ correspond to the *same* physical state. The observables for the system described by $\hat{P}$ must therefore Poisson-commute with all functions in Gau. Hence one might expect the Poisson algebra of physical observables to be given by the quotient $\mathcal{F}/\text{Gau}$. However, since Gau is not a Lie ideal in $\mathcal{F}$ the quotient is not a Lie algebra and hence not a Poisson algebra either. The way to proceed is to consider the biggest Poisson subalgebra of $\mathcal{F}$ which contains Gau as Lie ideal and then take the quotient. Hence we make the following

DEFINITION 6. The *Lie idealiser* of $\text{Gau} \subset \mathcal{F}$ is

$$\mathcal{I}_{\text{Gau}} := \{f \in \mathcal{F} \mid \{f, g\}|_{\hat{P}} = 0,\ \forall g \in \text{Gau}\}\,. \tag{85}$$

$\mathcal{I}_{\text{Gau}}$ is the space of functions which, in Dirac's terminology [4], are said to *weakly* commute with all gauge functions $g \in \text{Gau}$; that is, $\{f, g\}$ is required to vanish only *after* restriction to $\hat{P}$.

LEMMA 4. *$\mathcal{I}_{\text{Gau}}$ is a Poisson subalgebra of $\mathcal{F}$ which contains* Gau *as ideal.*

Proof. Let $f, g \in \mathcal{I}_{\text{Gau}}$ and $h \in \text{Gau}$. Then clearly $f + g \in \mathcal{I}_{\text{Gau}}$ and also $\{f \cdot g, h\}|_{\hat{P}} = f \cdot \{g, h\}|_{\hat{P}} + g \cdot \{f, h\}|_{\hat{P}} = 0$ (since each term vanishes), hence $\mathcal{I}_{\text{Gau}}$ is an associative subalgebra. Moreover, using the Jacobi identity, we have

$$\{\{f, g\}, h\}|_{\hat{P}} = \{\underbrace{\{h, g\}}_{\in\text{Gau}}, f\}|_{\hat{P}} + \{\underbrace{\{f, h\}}_{\in\text{Gau}}, g\}|_{\hat{P}} = 0\,, \tag{86}$$

which establishes that $\mathcal{I}_{\text{Gau}}$ is also a Lie subalgebra. Gau is obviously an associative ideal in $\mathcal{I}_{\text{Gau}}$ (since it is such an ideal in $\mathcal{F}$) and, by definition, also a Lie ideal. Hence it is a Poisson ideal. □

It follows from its very definition that $\mathcal{I}_{\text{Gau}}$ is maximal in the sense that there is no strictly larger subalgebra in $\mathcal{F}$ in which Gau is a Poisson algebra. Now we can define the algebra of *physical observables*:

Definition 7. The *Poisson algebra of physical observables* is given by

$$\mathcal{O}_{\text{phys}} := \mathcal{I}_{\text{Gau}}/\text{Gau}\,. \tag{87}$$

Since the restriction to $\hat{P}$ of a Hamiltonian vector field X_g is tangent to $\hat{P}$ if $g \in \text{Gau}$ (by Lemma 3 and coisotropy), we have

$$\begin{aligned} \mathcal{I}_{\text{Gau}} =& \{f \in \mathcal{F} \mid X_g(f)|_{\hat{P}} = 0, \forall g \in \text{Gau}\} \\ =& \{f \in \mathcal{F} \mid X_g|_{\hat{P}}(f|_{\hat{P}}) = 0, \forall g \in \text{Gau}\}\,, \end{aligned} \tag{88}$$

which shows that $\mathcal{I}_{\text{Gau}}$ is the subspace of all functions in $\mathcal{F}$ whose restrictions to $\hat{P}$ are constant on each connected leaf of the foliation tangent to the integrable subbundle $T^{\perp}\hat{P}$. If the space of leaves is a smooth manifold[9] it has a natural symplectic structure. In this case it is called the *reduced phase space* $(\bar{P}, \bar{\omega})$. $\mathcal{O}_{\text{phys}}$ can then be naturally identified with the Poisson algebra of (say C^∞-) functions thereon.

We finally mention that instead of the Lie idealiser $\mathcal{I}_{\text{Gau}}$ we could not have taken the Lie centraliser

$$\begin{aligned} \mathcal{C}_{\text{Gau}} :=& \{f \in \mathcal{F} \mid \{f, g\} = 0,\ \forall g \in \text{Gau}\} \\ =& \{f \in \mathcal{F} \mid X_g(f) = 0,\ \forall g \in \text{Gau}\}\,, \end{aligned} \tag{89}$$

which corresponds to the space of functions which, in Dirac's terminology [4], *strongly* commute with all gauge functions. This space is generally far too small, as can be seen from the following

Lemma 5. *If $\hat{P}$ is a closed subset of P we have*

$$\mathit{Span}\{X_g(p),\ g \in \text{Gau}\} = \begin{cases} T_p^{\perp}\hat{P} & \textit{for } p \in \hat{P} \\ T_pP & \textit{for } p \in P - \hat{P}\,. \end{cases} \tag{90}$$

Proof. For $p \in \hat{P}$ we know from Lemma 3 that $X_g(p) \in T_p^{\perp}\hat{P}$. Locally we can always find codim$(\hat{P})$ functions $g_i \in \text{Gau}$ whose differentials dg_i (and hence whose vector fields X_{g_i}) at p are linearly independent. To see that the $X_g(p)$ span all of T_pP for $p \notin \hat{P}$, we choose a neighbourhood U of p such that $U \cap \hat{P} = \emptyset$ (such U exists since $\hat{P} \subset P$ is closed by hypothesis) and $\beta \in C^\infty(P)$ such that $\beta|_U \equiv 1$ and $\beta|_{\hat{P}} \equiv 0$. Then $\beta \cdot h \in \text{Gau}$ for all $h \in C^\infty(P)$ and $(\beta \cdot h)|_U = h|_U$, which shows that $\text{Span}\{X_g(p), g \in \text{Gau}\} = \text{Span}\{X_g(p), g \in C^\infty(P)\} = T_pP$. □

This Lemma immediately implies that functions which strongly commute with all gauge functions must have altogether vanishing directional derivatives outside

[9] The 'space of leaves' is the quotient space with respect to the equivalence relation 'lying on the same leaf'. If the leaves are the orbits of a group action (the group of gauge transformations) then this quotient will be a smooth manifold if the group action is smooth, proper, and free (cf. Sect. 4.1 of [1]).

$\hat{P}$, that is, they must be constant on any connected set outside $\hat{P}$. By continuity they must be also constant on any connected subset of $\hat{P}$. Hence the condition of strong commutativity is far too restrictive.

Sometimes strong commutativity is required, but only with a somehow preferred subset ϕ_α, $\alpha = 1, \cdots, \mathrm{codim}(\hat{P})$, of functions in Gau; for example, the component functions of a momentum map (cf. Sect. 4.2 of [1]) of a group (the group of gauge transformations) that acts symplectomorphically (i.e. ω-preserving) on P. The size of the space of functions on P that strongly commute with all ϕ_α will generally depend delicately on the behaviour of the ϕ_α off the constraint surface, and may again turn out to be too small. The point being that even though the leaves generated by the ϕ_α may behave well *within* the zero-level set of all ϕ_α (the constraint surface), so that sufficiently many invariant (i.e. constant along the leaves) functions exist on the constraint surface, the leaves may become more 'wild' in infinitesimal neighbouring level sets, thereby forbidding most of these functions to be extended to some invariant functions in a neighbourhood of $\hat{P}$ in P. See Sect. 3 of [2] for an example and more discussion of this point.

Appendix 1: Geometry of Hamiltonian Systems

A *symplectic manifold* is a pair (P, ω), where P is a differentiable manifold and ω is a closed (i.e. $d\omega = 0$) 2-form which is non-degenerate (i.e. $\omega_p(X_p, Y_p) = 0$, $\forall X_p \in T_pP$, implies $Y_p = 0$ for all $p \in P$). The last condition implies that P is even dimensional. Let $C^\infty(P)$ denote the set of infinitely differentiable, real valued functions on P and $\mathcal{X}(P)$ the set of infinitely differentiable vector fields on P. $\mathcal{X}(P)$ is a real Lie algebra (cf. Appendix 2) whose Lie product is the commutator of vector fields. There is a map $X : C^\infty(P) \to \mathcal{X}(P)$, $f \mapsto X_f$, uniquely defined by[10]

$$X_f \vdash \omega = -df\,. \tag{91}$$

The kernel of X in $C^\infty(P)$ are the constant functions and the image of X in $\mathcal{X}(P)$ are called Hamiltonian vector fields. The Lie derivative of ω with respect to an Hamiltonian vector field is always zero:

$$L_{X_f}\omega = d(X_f \vdash \omega) = -ddf = 0\,, \tag{92}$$

where we used the following identity for the Lie derivative L_Z with respect to any vector field Z on forms of any degree:

$$L_Z = d \circ (Z \vdash) + (Z \vdash) \circ d\,. \tag{93}$$

The map X can be used to turn C^∞ into a Lie algebra. The Lie product $\{\cdot, \cdot\}$ on C^∞ is called Poisson bracket and defined by

$$\{f, g\} := \omega(X_f, X_g) = X_f(g) = -X_g(f)\,, \tag{94}$$

[10] For notation recall footnote 8.

where the 2nd and 3rd equality follows from (91). With respect to this structure the map X is a homomorphism of Lie algebras:

$$\begin{aligned} X_{\{f,g\}} \vdash \omega = -d\{f,g\} &\stackrel{94}{=} -d(X_g \vdash X_f \vdash \omega) \\ &\stackrel{93,91}{=} -L_{X_g}(X_f \vdash \omega) \\ &\stackrel{92}{=} [X_f, X_g] \vdash \omega\,. \end{aligned} \tag{95}$$

One may say that the map X has pulled back the Lie structure from $\mathcal{X}(P)$ to $C^\infty(P)$. Note that (95) also expresses the fact that Hamiltonian vector fields form a Lie subalgebra of $\mathcal{X}(P)$.

Special symplectic manifolds are the cotangent bundles. Let M be a manifold and $P = T^*M$ its cotangent bundle with projection $\pi : T^*M \to M$. On P there exists a naturally given 1-form field (i.e. section of $T^*P = T^*T^*M$), called the canonical 1-form (field) θ:

$$\theta_p := p \circ \pi_*|_p\,. \tag{96}$$

In words, application of θ to $Z_p \in T_pP$ is as follows: project Z_p by the differential π_*, evaluated at p, into T_xM, where $x = \pi(p)$, and then act upon it by p, where $p \in \pi^{-1}(x) = T^*_xM$ is understood as 1-form on M. The exterior differential of the canonical 1-form defines a symplectic structure on P (the minus sign being conventional):

$$\omega := -d\theta\,. \tag{97}$$

In canonical (Darboux-) coordinates ($\{q^i\}$ on M and $\{p_i\}$ on the fibres $\pi^{-1}(x)$) one has

$$\theta = p_i\, dq^i \quad \text{and} \quad \omega = dq^i \wedge dp_i\,, \tag{98}$$

so that

$$\{f,g\} = \sum_i \left(\frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \right)\,. \tag{99}$$

In this coordinates the Hamiltonian vector field X_f reads:

$$X_f = (\partial_{p_i} f)\partial_{q^i} - (\partial_{q^i} f)\partial_{p_i}\,. \tag{100}$$

It is important to note that Hamiltonian vector fields need not be complete, that is, their flow need not exist for all flow parameters $t \in \mathbb{R}$. For example, consider $P = \mathbb{R}^2$ in canonical coordinates. The flow map $\mathbb{R} \times P \to P$ is then given by $(t, (q_0, p_0)) \mapsto (q(t; q_0, p_0), p(t; q_0, p_0))$, where the functions on the right hand side follow through integration of $X_f = \dot{q}(t)\partial_q + \dot{p}(t)\partial_p$, i.e.

$$\dot{q}(t) = (\partial_p f)(q(t), p(t)) \quad \text{and} \quad \dot{p}(t) = -(\partial_q f)(q(t), p(t))\,, \tag{101}$$

with initial conditions $q(0) = q_0$, $p(0) = p_0$. As simple exercises one readily solves for the flows of $f(q,p) = h(q)$, $f(q,p) = h(p)$, where $h : P \to \mathbb{R}$ is some

C^1-function, or for the flow of $f(q,p) = qp$. All these are complete. But already for $f(q,p) = q^2 p$ we obtain

$$q(t; q_0, p_0) = \frac{q_0}{1 - q_0 t} \quad \text{and} \quad p(t; q_0, p_0) = p_0 \, (1 - q_0 t)^2 \,, \tag{102}$$

which (starting from $t = 0$) exists only for $t < 1/q_0$ when $q_0 > 0$ and only for $t > 1/q_0$ when $q_0 < 0$.

Appendix 2: The Lie Algebra of $sl(2, \mathbb{R})$ and the Absence of Non-trivial, Finite-Dimensional Representations by Anti-unitary Matrices

Let us first recall the definition of a Lie algebra:

DEFINITION 8. A *Lie algebra over* $\mathbb{F}$ (here standing for $\mathbb{R}$ or $\mathbb{C}$) is a vector-space, L, over $\mathbb{F}$ together with a map $V \times V \rightarrow V$, called *Lie bracket* and denoted by $[\cdot,\cdot]$, such that the following conditions hold for all $X, Y, Z \in L$ and $a \in \mathbb{F}$:

$$[X, Y] = -[Y, X] \qquad \text{antisymmetry}\,, \tag{103}$$

$$[X, Y + aZ] = [X, Y] + a[X, Z] \qquad \text{linearity}\,, \tag{104}$$

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \qquad \text{Jacobi identity}\,. \tag{105}$$

Note that (103) and (104) together imply linearity also in the first entry. Any associative algebra (with multiplication '$\cdot$') is automatically a Lie algebra by defining the Lie bracket to be the commutator $[X, Y] := X \cdot Y - Y \cdot X$ (associativity then implies the Jacobi identity). Important examples are Lie algebras of square matrices, whose associative product is just matrix multiplication.

A sub vector-space $L' \subseteq L$ is a *sub Lie-algebra*, iff $[X, Y] \in L'$ for all $X, Y \in L'$. A sub Lie-algebra is an *ideal*, iff $[X, Y] \in L'$ for all $X \in L'$ and all $Y \in L$ (sic!). Two ideals always exist: L itself and $\{0\}$; they are called the trivial ideals. A Lie algebra is called *simple*, iff it contains only the trivial ideals. A map $\phi : L \rightarrow L'$ between Lie algebras is a *Lie homomorphism*, iff it is linear and satisfies $\phi([X, Y]) = [\phi(X), \phi(Y)]$ for all $X, Y \in L$. Note that we committed some abuse of notation by denoting the (different) Lie brackets in L and L' by the same symbol $[\cdot,\cdot]$. The kernel of a Lie homomorphism ϕ is defined by $\mathrm{kernel}(\phi) := \{X \in L \mid \phi(X) = 0\}$ and obviously an ideal in L.

The Lie algebra denoted by $sl(2, \mathbb{F})$ is defined by the vector space of traceless 2×2 - matrices with entries in $\mathbb{F}$. A basis is given by

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{106}$$

Its commutation relations are:

$$[H, E_+] = 2E_+ \,, \tag{107}$$

$$[H, E_-] = -2E_- \,, \tag{108}$$

$$[E_+, E_-] = H \,. \tag{109}$$

The first thing we prove is that $sl(2,\mathbb{F})$ is simple. For this, suppose $X = aE_+ + bE_- + cH$ is a member of an ideal $I \subseteq sl(2,\mathbb{F})$. From (107-109) we calculate

$$[E_+,[E_+,X]] = -2bE_+\,, \tag{110}$$

$$[E_-,[E_-,X]] = -2aE_-\,. \tag{111}$$

Suppose first $b \neq 0$, then (110) shows that $E_+ \in I$. Then (109) implies $H \in I$, which in turn implies through (108) that $E_- \in I$ and hence that $I = L$. Similarly one concludes for $a \neq 0$ that $I = L$. Finally assume $a = b = 0$ and $c \neq 0$ so that $H \in I$. Then (107) and (108) show that E_+ and E_- are in I, so again $I = L$. Hence we have shown that $I = L$ or $I = \{0\}$ are the only ideals.

Next consider the Lie algebra $u(n)$ of anti-Hermitean $n \times n$ matrices. It is the Lie algebra of the group $U(n)$ of unitary $n \times n$ matrices. If the group $SL(2,\mathbb{F})$ had a finite-dimensional unitary representation, i.e. if a group homomorphism $D : SL(2,\mathbb{F}) \to U(n)$ existed for some n, then we would also have a Lie homomorphism $D_* : sl(2,\mathbb{F}) \to u(n)$ by simply taking the derivative of the map D at e (= identity of $SL(2,\mathbb{R})$). We will now show that, for any integer $n \geq 1$, any Lie homomorphism $\phi : sl(2,\mathbb{F}) \to u(n)$ is necessarily the constant map onto $0 \in u(n)$. In other words, non-trivial Lie homomorphism from $sl(2,\mathbb{F})$ to $u(n)$ do not exist. On the level of groups this implies that non-trivial (i.e. not mapping everything into the identity), finite dimensional, unitary representations of $SL(2,\mathbb{F})$ do not exist. Note that for $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$ these are (the double covers of) the proper orthochronous Lorentz groups in 2+1 and 3+1 dimensions respectively.

To prove this result, assume $T : sl(2,\mathbb{F}) \to u(n)$ is a Lie homomorphism. To save notation we write $T(H) =: A$ and $T(E_\pm) =: B_\pm$. Since T is a Lie homomorphism we have $[A,B_+] = 2B_+$, which implies

$$\mathrm{trace}(B_+^2) = \tfrac{1}{2}\mathrm{trace}\left(B_+\left(AB_+ - B_+A\right)\right) = 0\,, \tag{112}$$

where in the last step we used the cyclic property of the trace. But B_+ is anti-Hermitean, hence diagonalisable with purely imaginary eigenvalues $\{\mathrm{i}\lambda_1, \cdots, \mathrm{i}\lambda_n\}$ with $\lambda_i \in \mathbb{R}$. The trace on the left side of (112) is then $-\sum_i \lambda_i^2$, which is zero iff $\lambda_i = 0$ for all i, i.e. iff $B_+ = 0$. Hence $E_+ \in \mathrm{kernel}(T)$, which in turn implies $\mathrm{kernel}(T) = sl(2,\mathbb{F})$ since the kernel – being an ideal – is either $\{0\}$ or all of $sl(2,\mathbb{F})$ by simplicity. This proves the claim, which is stated as Lemma 1 of the main text

References

1. Ralph Abraham and Jerrold E. Marsden: *Foundations of Mechanics*, second edition (The Benjamin/Cummings Publishing Company, Reading, Massachusetts, 1978)
2. Martin Bordemann, Hans-Christian Herbig, and Stefan Waldmann: BRST cohomology and phase space reduction in deformation quantisation. *Communications of Mathematical Physics* **210**, 107–144 (2000). math.QA/9901015

3. Paul Dirac: *The Principles of Quantum Mechanics*, fourth edition. The International Series of Monographs in Physics 27 (Oxford University Press, Oxford (UK), 1958), 1982 reprint of fourth edition
4. Paul Dirac: *Lectures on Quantum Mechanics.* Belfer Graduate School of Science Monographs Series, Number Two (Yeshiva University, New York, 1964)
5. Mark Gotay: Functorial geometric quantization and van Hove's theorem. *Int. J. Theor. Phys.* **19**, 139–161 (1980)
6. Mark Gotay: On a full quantization of the torus. In J.-P. Antoine et al., editors, *Quantization, Coherent States, and Complex Structure*, pages 55–62, New York, 1995. Białowieża, 1994, Plenum Press math-ph/9507005.
7. Mark Gotay: A Groenewold - van Howe theorem for S^2. *Trans. Am. Math. Soc.* **348**, 1579–1597 (1996). math-ph/9502008
8. Mark Gotay: On the Groenewold - van Howe problem for $\mathbb{R}^{2n}$. *J. Math. Phys.* **40**, 2107–2116 (1999). math-ph/9809015
9. Hip Groenewold: On the principles of elementary quantum mechanics. *Physica* **12**, 405–460 (1946)
10. Rudolf Haag: *Local Quantum Physics: Fields, Particles, Algebras*, second edition. Texts and Monographs in Physics (Springer Verlag, Berlin, 1996)
11. Marc Henneaux and Claudio Teitelboim: *Quantization of Gauge Systems*, first edition (Princeton University Press, Princeton, USA, 1992)
12. Chris Isham: Topological and global aspects of quantum theory. In B.S. DeWitt and R. Stora, editors, *Relativity, Groups and Topology II*, pages 1059–1290, Les Houches 1983, Session XL (North-Holland Physics Publishing, Amsterdam, 1984)
13. Erich Joos, H.-Dieter Zeh, Claus Kiefer, Domenico Giulini, Joachim Kupsch, and Ion-Olimpiu Stamatescu: *Decoherence and the Appearence of a Classical World in Quantum Theory*, second edition (Springer Verlag, Berlin, 2003)
14. Mike Reed and Barry Simon: *Functional Analysis* Methods of Modern Mathematical Physics, first edition (Academic Press, New York, 1972)
15. Léon van Howe: Sur certaines représentations unitaires d'un groupe infini de transformations. *Mem. Acad. Roy. Belg. (Cl. Sci.)* **26**, 61–102 (1951)
16. Léon van Howe: Sur le problème des relations entre les transformations unitaires de la mécanique quantique et les transformations canoniques de la mécanique classique. *Acad. Roy. Belg. Bull. (Cl. Sci.)* **37**, 610–620 (1951)

Lectures on Loop Quantum Gravity

Thomas Thiemann

MPI für Gravitationsphysik, Albert-Einstein-Institut, Am Mühlenberg 1,
14476 Golm, Germany

Abstract. Quantum General Relativity (QGR), sometimes called Loop Quantum Gravity, has matured over the past fifteen years to a mathematically rigorous candidate quantum field theory of the gravitational field. The features that distinguish it from other quantum gravity theories are 1) *background independence* and 2) *minimality of structures.*

Background independence means that this is a non-perturbative approach in which one does not perturb around a given, distinguished, classical background metric, rather arbitrary fluctuations are allowed, thus precisely encoding the quantum version of Einstein's radical perception that *gravity is geometry.*

Minimality here means that one explores the logical consequences of bringing together the two fundamental principles of modern physics, namely general covariance and quantum theory, without adding any experimentally unverified additional structures such as extra dimensions, extra symmetries or extra particle content beyond the standard model. While this is a very conservative approach and thus maybe not very attractive to many researchers, it has the advantage that pushing the theory to its logical frontiers will undoubtedly either result in a successful theory or derive exactly which extra structures are required, if necessary. Or put even more radically, it may show which basic principles of physics have to be given up and must be replaced by more fundamental ones.

QGR therefore is, by definition, not a unified theory of all interactions in the standard sense, since such a theory would require a new symmetry principle. However, it unifies all presently known interactions in a new sense by quantum mechanically implementing their common symmetry group, the four-dimensional diffeomorphism group, which is almost completely broken in perturbative approaches.

In this contribution we summarize the present status of Canonical Quantum General Relativity (QGR), also known as "Loop Quantum Gravity". Our presentation tries to be precise and at the same time technically not too complicated by skipping the proofs of all the statements made. These many missing details, which are relevant to the serious reader, can be found in the notation used in the recent, close to exhaustive review [1] and references therein. Of course, in order to be useful as a text for first reading we did not include all the relevant references here. We apologize for that to the researchers in the field but we hope that a close to complete list of their work can be found in [1]. Nice reports, treating complementary subjects of the field and more general aspects of quantum gravity can be found in [2].

The text is supplemented by numerous exercises of varying degree of difficulty whose purpose is to cut the length of the exposition and to arouse interest

T. Thiemann, Lectures on Loop Quantum Gravity, Lect. Notes Phys. **631**, 41–135 (2003)
http://www.springerlink.com/

in further studies. Solving the problems is not at all mandatory for an understanding of the material, however, the exercises contain further information and thus should be looked at even on a first reading.

On the other hand, if one solves the problems then one should get a fairly good insight into the techniques that are important in QGR and in principle could serve as a preparation for a diploma thesis or a dissertation in this field. The problems sometimes involve mathematics that may be unfamiliar to students, however, this should not scare off but rather encourage the serious reader to learn the necessary mathematical background material. Here is a small list of mathematical texts, from the author's own favourites, geared at theoretical and mathematical physicists, that might be helpful:

- *General*
  A fairly good encyclopedia is
  Y. Choquet-Bruhat, C. DeWitt-Morette, "Analysis, Manifolds and Physics", North Holland, Amsterdam, 1989
- *General Topology*
  A nice text, adopting almost no prior knowledge is
  J.R. Munkres, "Toplogy: A First Course", Prentice Hall Inc., Englewood Cliffs (NJ), 1980
- *Differential and Algebraic Geometry*
  A modern exposition of this classical material can be found in
  M. Nakahara, "Geometry, Topology and Physics", Institute of Physics Publishing, Bristol, 1998
- *Functional Analysis*
  The number one, unbeatable and close to complete exposition is
  M. Reed, B. Simon, "Methods of Modern Mathematical Physics", vol. 1–4, Academic Press, New York, 1978
  especially volumes one and two.
- *Measure Theory*
  An elementary introduction to measure theory can be found in the beautiful book
  W. Rudin, "Real and Complex Analysis", McGraw-Hill, New York, 1987
- *Operator Algebras*
  Although we do not really make use of C^*-algebras in this review, we hint at the importance of the subject, so let us include
  O. Bratteli, D.W. Robinson, "Operator Algebras and Quantum Statistical Mechanics", vol. 1,2, Springer Verlag, Berlin, 1997
- *Harmonic Analysis on Groups*
  Although a bit old, it still contains a nice collection of the material around the Peter & Weyl theorem:
  N.J. Vilenkin, "Special Functions and the Theory of Group Representations", American Mathematical Society, Providence, Rhode Island, 1968
- *Mathematical General Relativity*
  The two leading texts on this subject are
  R.M. Wald, "General Relativity", The University of Chicago Press, Chicago, 1989

S. Hawking, Ellis, "The Large Scale Structure of Spacetime", Cambridge University Press, Cambridge, 1989

- *Mathematical and Physical Foundations of Ordinary QFT*
  The most popular books on axiomatic, algebraic and constructive quantum field theory are
  R.F. Streater, A.S. Wightman, "PCT, Spin and Statistics, and all that", Benjamin, New York, 1964
  R. Haag, "Local Quantum Physics", 2nd ed., Springer Verlag, Berlin, 1996
  J. Glimm, A. Jaffe, "Quantum Physics", Springer-Verlag, New York, 1987

In the first part we motivate the particular approach to a quantum theory of gravity, called (Canonical) Quantum General Relativity, and develop the classical foundations of the theory as well as the goals of the quantization programme.

In the second part we list the solid results that have been obtained so far within QGR. Thus, we will apply step by step the quantization programme outlined at the end of Sect. 1.3 to the classical theory that we defined in Sect. 1.2. Up to now, these steps have been completed approximately until step vii) at least with respect to the Gauss- and the spatial diffeomorphism constraint. The analysis of the Hamiltonian constraint has also reached level vii) already, however, its classical limit is presently under little control which is why we discuss it in part three where current research topics are listed.

In the third part we discuss a selected number of active research areas. The topics that we will describe already have produced a large number of promising results, however, the analysis is in most cases not even close to being complete and therefore the results are less robust than those that we have obtained in the previous part.

Finally, in the fourth part we summarize and list the most important open problems that we faced during the discussion in this report.

1 Motivation and Introduction

1.1 Motivation

Why Quantum Gravity in the 21'st Century? Students that plan to get involved in quantum gravity research should be aware of the fact that in our days, when financial resources for fundamental research are more and more cut and/or more and more absorbed by research that leads to practical applications on short time scales, one should have a good justification for why tax payers should support any quantum gravity research at all. This seems to be difficult at first due to the fact that even at CERN's LHC we will be able to reach energies of at most 10^4 GeV which is **fifteen orders of magnitude below the Planck scale** which is the energy scale at which quantum gravity is believed to become important. Therefore one could argue that quantum gravity research in the 21'st century is of purely academic interest only.

To be sure, it is a shame that one has to justify fundamental research at all, a situation unheard of in the beginning of the 20'th century which probably was part of the reason for why so many breakthroughs especially in fundamental physics have happened in that time. Fundamental research can work only in absence of any pressure to produce (mainstream) results, otherwise new, radical and independent thoughts are no longer produced. To see the time scale on which fundamental research leads to practical results, one has to be aware that General Relativity (GR) and Quantum Theory (QT) were discovered in the 20's and 30's already but it took some 70 years before quantum mechanics through, e.g. computers, mobile phones, the internet, electronic devices or general relativity through e.g. space travel or the global positioning system (GPS) became an integral part of life of a large fraction of the human population. Where would we be today if the independent thinkers of those times were forced to do practical physics due to lack of funding for analyzing their fundamental questions?

Of course, in the beginning of the 20'th century, one could say that physics had come to some sort of crisis, so that there was urgent need for some revision of the fundamental concepts: Classical Newtonian mechanics, classical electrodynamics and thermodynamics were so well understood that Max Planck himself was advised not to study physics but engineering. However, although from a practical point of view all seemed well, there were subtle inconsistencies among these theories if one drove them to their logical frontiers. We mention only three of them:

1. Although the existence of atoms was by far not widely accepted at the end of the 19th century (even Max Planck denied them), if they existed then there was a serious flaw, namely, how should atoms be stable? Accelerated charges radiate Bremsstrahlung according to Maxwell's theory, thus an electron should fall into the nucleus after a finite amount of time.
2. If Newton's theory of absolute space and time was correct then the speed of light should depend on the speed of the inertial observer. The fact that such velocity dependence was ruled out to quadratic order in v/c in the famous Michelson-Morley experiment was explained by postulating an unknown medium, called ether, with increasingly (as measurement precision was refined) bizarre properties in order to conspire to a negative outcome of the interferometer experiment and to preserve Newton's notion of space and time.
3. The precession of mercury around the sun contradicted the ellipses that were predicted by Newton's theory of gravitation.

Today we easily resolve these problems by 1) quantum mechanics, 2) special relativity and 3) general relativity. Quantum mechanics does not allow for continuous radiation but predicts a discrete energy spectrum of the atom, special relativity removed the absolute notion of space and time and general relativity generalizes the static Minkowski metric underlying special relativity to a dynamical theory of a metric field which revolutionizes our understanding of gravity not as a force but as geometry. Geometry is curved at each point ithiem11n a manifold proportional to the matter density at that point and in turn curvature tells matter what are the straightest lines (geodesics) along which to move. The ether became completely unnecessary by changing the foundation of physics and beautifully demonstrates that driving a theory to its logical frontiers can make extra structures redundant, what one had to change is the basic principles of physics.[1]

This historic digression brings us back to the motivation for studying quantum gravity in the beginning of the 21st century. The question is whether fundamental physics also today is in a kind of crisis. We will argue that indeed we are in a situation not dissimilar to that of the beginning of the 20th century: Today we have very successful theories of all interactions. Gravitation is described by general relativity, matter interactions by the standard model of elementary particle physics. As classical theories, their dynamics is summarized in the classical Einstein equations. However, there are several problems with these theories, some of which we list below:

i) *Classical-Quantum Inconsistency*
The fundamental principles collide in the classical Einstein equations

$$\boxed{\underbrace{R_{\mu\nu} \;-\; \tfrac{1}{2}\, R\, g_{\mu\nu}}_{\text{Geometry (GR, gen. covariance)}} \;=\; \underbrace{\kappa\, T_{\mu\nu}(g)}_{\text{Matter (Stand.model, QT)}}}$$

These equations relate matter density in form of the energy momentum tensor $T_{\mu\nu}$ and geometry in form of the Ricci curvature tensor $R_{\mu\nu}$. Notice that the metric tensor $g_{\mu\nu}$ enters also the definition of the energy momentum tensor. However, while the left hand side is described until today only by a classical theory, the right hand side is governed by a quantum field theory (QFT). Since complex valued functions and operators on a Hilbert space are two completely different mathematical objects, the only way to make sense out of the above equations while keeping the classical and quantum nature of geometry and matter respectively is to take expectation values of the right hand side, that is,

$$R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu} \;=\; \kappa \;< \hat{T}_{\mu\nu}(g_0) >, \quad \kappa = \frac{8\pi G_{\mathrm{Newton}}}{c^4}$$

Here g_0 is an arbitrary background metric, say the Minkowski metric $\eta = \mathrm{diag}(-1,1,1,1)$. However, even if the state with respect to which the expectation value is taken is the vacuum state ψ_{g_0} with respect to g_0 (the notion of vacuum depends on the background metric itself, see below), the right hand side is generically non-vanishing due to the vacuum fluctuations, enforcing $g = g_1 \neq g_0$. Hence, in order to make this system of equations consistent, one could iterate the procedure by computing the vacuum state ψ_{g_1} and reinserting g_1 into $\hat{T}_{\mu\nu}(.)$, resulting in $g_2 \neq g_1$ etc. hoping that the procedure converges. However, this is generically not the case and results in "run-away solutions" [3].

[1] Notice, that the stability of atoms is still not satisfactorily understood even today because the full problem also treats the radiation field, the nucleus and the electron as quantum objects which ultimately results in a problem in QED, QFD and QCD for which we have no entirely satisfactory description today, see below.

Hence, we are enforced to quantize the metric itself, that is, we need a quantum theory of gravity resulting in the

$$\boxed{\begin{array}{c}\widehat{R}_{\mu\nu} - \frac{1}{2}\widehat{g}_{\mu\nu}\,\widehat{R} = \underline{\kappa\,\widehat{T}_{\mu\nu}(\widehat{g})} \\ \text{"}\underline{\text{Quantum}}\text{-Einstein-Equations"}\end{array}} \tag{1}$$

The inverted commas in this equation are to indicate that this equation is to be made rigorous in a Hilbert space context. QGR is designed to exactly do that, see Sect. 3.1.

ii) *General Relativity Inconsistencies*
It is well-known that classical general relativity is an incomplete theory because it predicts the existence of so-called spacetime singularities, regions in spacetime where the curvature or equivalently the matter density becomes infinite [4]. The most prominent singularities of this kind are black hole and big bang singularities and such singularities are generic as shown in the singularity theorems due to Hawking and Penrose. When a singularity appears it means that the theory has been pushed beyond its domain of validity, certainly when matter collapses it reaches a state of extreme energy density at which quantum effects become important. A quantum theory of gravity could be able to avoid these singularities in a similar way as quantum mechanics explains the stability of atoms. We will see that QGR is able to achieve this, at least in the simplified context of "Loop Quantum Cosmology", see Sect. 3.2.

iii) *Quantum Field Theory Inconsistencies*
As is well-known, QFT is plagued by UV (or short distance) divergences. The fundamental operators of the theory are actually not operators but rather operator-valued distributions and usually interesting objects of the theory are (integrals of) polynomials of those evaluated in the same point. However, the product of distributions is, generally, ill-defined. The appearance of these divergences is therefore, on the one hand, not surprising, on the other hand it indicates again that the theory is incomplete: In a complete theory there is no room for infinities. Thus, either the appropriate mathematical framework has not been found yet, or they arise because one neglected the interaction with the gravitational field. In fact, in renormalizable theories one can deal with these infinities by renormalization, that is, one introduces a short distance cut-off (e.g. by point splitting the operator-valued distributions) and then redefines masses and coupling constants of the theory in a cut-off dependent way such that they stay finite as the cut-off is sent to zero. This redefinition is done in the framework of perturbation theory (Feynman diagrammes) by subtracting counter terms from the original Lagrangean which are formally infinite and a theory is said to be renormalizable if the number of algebraically different such counter terms is finite.
The occurrence of UV singularities is in deep conflict with general relativity due to the following reason: In perturbation theory, the divergences have

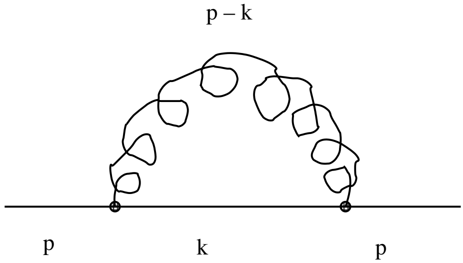


Fig. 1. One loop correction to the electron propagator in QED

their origin in Feynman loop integrals in momentum space where the inner loop 4-momentum $k = (E, P)$ can become arbitrarily large, see Fig. 1 for an example from QED (mass renormalization). Now such virtual (off-shell) particles with energy E and momentum P have a spatial extension of the order of the Compton radius $\lambda = \hbar/P$ and a mass of the order of E/c^2. Classical general relativity predicts that this lump of energy turns into a black hole once λ reaches the Schwarzschild radius of the order of $r = GE/c^4$. In a Lorentz frame where $E \approx Pc$ this occurs at the Planck energy $E = E_{\mathrm{P}} = \sqrt{\hbar/\kappa c} \approx 10^{19}$ GeV or at the Planck length Compton radius $\ell_{\mathrm{P}} = \sqrt{\hbar\kappa} \approx 10^{-33}$ cm. However, when a (virtual) particle turns into a black hole it completely changes its properties. For instance, if the virtual particle is an electron then it is able to interact only electroweakly and thus can radiate only particles of the electrowak theory. However, once a black hole has formed, also Hawking processes are possible and now any kind of particles can be emitted, but at a different production rate. Of course, this is again an energy regime at which quantum gravity must be important and these qualitative pictures must be fundamentally wrong, however, they show that there is a problem with integrating virtual loops into the UV regime. In fact, these qualitative thoughts suggest that gravity could serve as a *natural cut-off* because a black hole of Planck length ℓ_{P} should decay within a Planck time unit $t_{\mathrm{P}} = \ell_{\mathrm{P}}/c \approx 10^{-43}$ s so that one has to integrate P only until E_{P}/c. Moreover, it indicates that spacetime geometry itself acquires possibly a *discrete structure* since arguments of this kind make it plausible that it is impossible to resolve spacetime distances smaller than ℓ_{P}, basically because the spacetime behind an event horizon is in some sense "invisible". These are, of course, only hopes and must be demonstrated within a concrete theory. We will see that QGR is able to precisely do that and its fundamental discreteness is in particular responsible for why the Bekenstein Hawking entropy of black holes is finite, see Sects. 2.2, 3.1 and 3.4.

So we see that there is indeed a fundamental inconsistency within the current description of fundamental physics comparable to the time before the discovery of GR and QT and its resolution, Quantum Gravity, will revolutionize not only our understanding of nature but will also drive new kinds of technology that we do not even dare to dream of today.

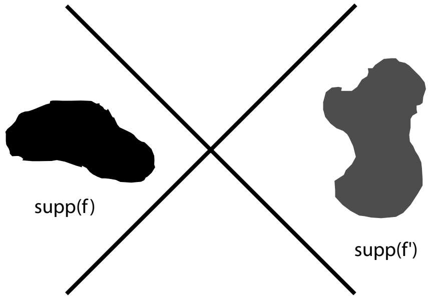


Fig. 2. Spacelike separated regions in Minkowski space

The Role of Background Independence. Given the fact that both QT and GR were discovered already more than 70 years ago and that people have certainly thought about quantizing GR since then and that matter interactions are more or less successfully described by ordinary quantum field theories (QFT), it is somewhat surprising that we do not yet have a quantum gravity theory. Why is it so much harder to combine gravity with the principles of quantum mechanics than for the other interactions? The short answer is that

Ordinary QFT only incorporates Special Relativity.

To see why, we just have to remember that ordinary QFT has an axiomatic definition, here for a scalar field for simplicity:
WIGHTMAN AXIOMS (Scalar Fields on Minkowski Space)

W1 *Poincaré Group* $\mathcal{P}$:
$\exists$ continuous, unitary representation $\hat{U}$ of $\mathcal{P}$ on a Hilbert space $\mathcal{H}$.

W2 Forward Lightcone *Spektral Condition*:
For the generators $\hat{P}^\mu$ of the translation subgroup of $\mathcal{P}$ holds $\eta_{\mu\nu}\hat{P}^\mu\hat{P}^\nu \leq 0, \;\; \hat{P}^0 \geq 0$.

W3 Existence and Uniqueness of a $\mathcal{P}$*-invariant Vacuum* Ω:
$\hat{U}(p)\Omega = \Omega \;\; \forall p \in \mathcal{P}$.

W4 $\mathcal{P}$*-Covariance*:

$$\hat{\phi}(f) := \int d^{D+1}x f(x)\hat{\phi}(x), \;\; f \in \mathcal{S}(R^{D+1})$$

$$\hat{\phi}(f_1)..\hat{\phi}(f_n)\Omega \text{ dense in } \mathcal{H} \text{ and } \hat{U}(p)\hat{\phi}(f)\hat{U}(p)^{-1} = \hat{\phi}(f \circ p).$$

W5 *Locality (Causality):*
If supp(f), supp(f') *spacelike separated* (see Fig. 2), then $[\hat{\phi}(f), \hat{\phi}(f')] = 0$.

It is obvious that due to the presence of the Minkowski background metric η we have available a large amount of structure which forms the fundament on which ordinary QFT is built. Roughly, we have the following scheme:

$$\begin{array}{ccccc}
\begin{array}{c}\eta_{\mu\nu}\\ \text{b.-metric}\end{array} & \Longrightarrow & \begin{array}{c}\mathcal{P}\\ \text{symm.-group}\end{array} & \Longrightarrow & \begin{array}{c}\hat{H}=\hat{P}^0\\ \text{Ham.-operator}\end{array}\\
\Downarrow & & & & \Downarrow\\
\begin{array}{c}(x-y)^2=0\\ \text{lightcone}\end{array} & \Longrightarrow & \begin{array}{c}\text{commutation}\\ \text{relations}\end{array} & \Longrightarrow & \begin{array}{c}\Omega\\ \text{ground state}\end{array}
\end{array}$$

Notice that a generic background metric has no symmetry group at all so that it is not straightforward to generalize these axioms to QFT on general curved backgrounds, however, since any metric is pointwise diffeomorphic to the Minkowski metric, a local generalization is possible and results in the so-called microlocal analysis in which the role of vacuum states is played by Hadamard states, see e.g. [5].

> The fundamental, radically new feature of Einstein's theory is that there is no background metric at all: The theory is background independent. The lightcones themselves are fluctuating, causality and locality become empty notions. The dome of ordinary QFT completely collapses.

Of course, there must be a regime in any quantum gravity theory where the quantum fluctuations of the metric operator are so tiny that we recover the well established theory of free ordinary quantum fields on a given background metric, however, the large fluctuations of the metric operator can no longer be ignored in extreme astrophysical or cosmological situations, such as near a black hole or big bang singularity.

People have tried to rescue the framework of ordinary QFT by splitting the metric into a background piece and a fluctuation piece

$$\begin{array}{ccccc}
g_{\mu\nu} & = & \eta_{\mu\nu} & + & h_{\mu\nu}\\
\uparrow & & \uparrow & & \uparrow\\
\text{full metric} & & \text{background (Minkowski)} & & \text{perturbation (graviton)}
\end{array} \tag{2}$$

which results in a Lagrangean for the graviton field $h_{\mu\nu}$ and could in principle be the definition of a graviton QFT on Minkowski space. However, there are serious drawbacks:

i) *Non-renormalizability*
The resulting theory is perturbatively non-renormalizable [6] as could have been expected from the fact that the coupling constant of the theory, the Planck area ℓ_{P}^2, has negative mass dimension (in Planck units). Even the supersymmetric extension of the theory, in any possible dimension has this bad feature [7]. It could be that the theory is non-perturbatively renormalizable, meaning that it has a non-Gaussian fix point in the language of Wilson, a possibility that has recently regained interest [8].

ii) *Violation of Background Independence*
The split of the metric performed above again distinguishes the Minkowski metric among all others and reintroduces therefore a background dependence. This violates the **key feature** of Einstein's theory and thus somehow does not sound correct, we better keep background independence if we want to understand how quantum mechanics can possibly work together with general covariance.

iii) *Violation of Diffeomorphism Covariance*
The split of the metric performed above is certainly not diffeomorphism covariant, it breaks the diffeomorphism group down to Poincaré group. Violation of fundamental, local gauge symmetries is usually considered as a bad feature in Yang-Mills theories on which all the other interactions are based, thus already from this point of view perturbation theory looks dangerous. As a side remark we see that background dependence and violation of general covariance are synonymous.

iv) *Gravitons and Geometry*
Somehow the whole idea of the gravitational interaction as a result of graviton exchange on a background metric contradicts Einstein's original and fundamental idea that gravity is geometry and not a force in the usual sense. Therefore such a perturbative description of the theory is very unnatural from the outset and can have at most a semi-classical meaning when the metric fluctuations are very tiny.

v) *Gravitons and Dynamics*
All that classical general relativity is about is how a metric evolves in time in an interplay with the matter present. It is clear that an initially (almost) Minkowskian metric can evolve to something that is far from Minkowskian at other times, an example being cosmological big bang situations or the collapse of initially diluted matter (evolved backwards). In such situations the assumption being made in (2), namely that h is "small" as compared to η is just not dynamically stable. In some sense it is like trying to use Cartesian coordinates for a sphere which can work at most locally.

All these points just naturally ask for a non-perturbative approach to quantum gravity. This, in turn, could also cure another unpleasant feature about ordinary QFT: Today we do not have a single example of a rigorously defined interacting ordinary QFT in four dimensions, in other words, the renormalizable theories that we have are only defined order by order in perturbation theory but the perturbation series diverges. A non-perturbative definition, to which we seem to be forced when coupling gravity anyway, might change this unsatisfactory situation.

It should be noted here that there is in fact a consistent perturbative description of a candidate quantum gravity theory, called string theory (or M-Theory nowadays) [9].[2] However, in order to achieve this celebrated rather non-trivial result, expectedly one must introduce extra structure: The theory lives in 10

[2] String theory is an ordinary QFT but not in the usual sense: It is an ordinary scalar QFT on a 2d Minkowski space, however, the scalar fields themselves are coordinates

(or 11) rather than 4 dimensions, it is necessarily supersymmetric and it has an infinite number of extra particles besides those that are needed to make the theory compatible with the standard model. Moreover, at least as presently understood, again the fundamental new ingredient of Einstein's theory, background independence, is violated in string theory. This current background dependence of string theory is supposed to be overcome once M-Theory has been rigorously defined.

At present only string theory has a chance to explain the matter content of our universe. The unification of symmetries is a strong guiding principle in physics as well and has been pushed also by Einstein in his programme of geometrization of physics attempting to unify electromagnetism and gravity in a five-dimensional Kaluza–Klein theory. The unification of the electromagnetic and the weak force in the electroweak theory is a prime example for the success of such ideas. However, unification of forces is an additional principle completely independent of background independence and is not necessarily what a *quantum theory* of gravity must achieve: Unification of forces can be analyzed at the purely *classical level*[3]. Thus, the only question is whether the theory can be quantized before unification or not (should unification of geometry and matter be realized in nature at all).

We are therefore again in a situation, similar to that before the discovery of special relativity, where we have the choice between a) preserving an old principle, here renormalizability of perturbative QFT on background spacetimes (M, η), at the price of introducing extra structure (extra unification symmetry), or b) replacing the old principle by a new principle, here non-perturbative QFT on a differentiable manifold M, without new hypothetical structure. At this point it unclear which methodology has more chances for success, historically there is evidence for either of them (e.g. the unification of electromagnetism and the massive Fermi model is evidence for the former, the replacement of Newton's notion of spacetime by special relativity is evidence for the latter) and it is quite possible that we actually need both ideas. In QGR we take the latter point of view to begin with since there maybe zillions of ways to unify forces and it is hard to judge whether there is a "natural one", therefore the approach is

of the ambient target Minkowski space which in this case is 10 dimensional. Thus, it is similar to a first quantized theory of point particles. The theory is renormalizable and presumably even finite order by order in perturbation theory but the perturbation series does not converge.

[3] In fact, e.g. the unified electroweak $SU(2)_L \times U(1)$ theory with its massless gauge bosons can be perfectly described by a classical Lagrangean. The symmetry broken, massive $U(1)$ theory can be derived from it, also classically, by introducing a constant background Higgs field (Higgs mechanism) and expanding the symmetric Lagrangean around it. It is true that the search for a massless, symmetric theory was inspired by the fact that a theory with massive gauge bosons is not renormalizable (so the motivation comes from quantum theory) and, given the non-renormalizability of general relativity, many take this as an indication that one must unify gravity with matter, one incarnation of which is string theory. However, the argument obviously fails should it be possible to quantize gravity *non-perturbatively*.

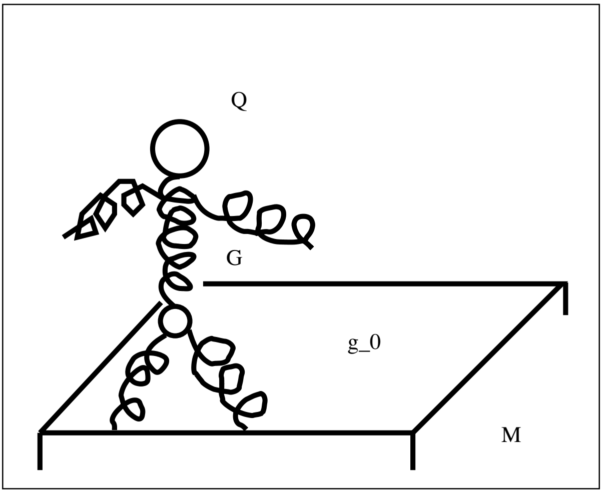


Fig. 3. QFT on Background Spacetime (M, g_0): Actor = Matter; Stage = Geometry + Manifold M

purposely conservative because we actually may be able to *derive* a natural way of unification, if necessary, if we drive the theory to its logical frontiers. Among the various non-perturbative approaches available we will choose the canonical one.

Pictorially, one could illustrate the deep difference between a background dependent QFT and background independent QFT as follows: In Fig. 3 we see matter in the form of QCD (notice the quark (Q) propagators, the quark-gluon vertices and the three- and four point gluon (G) vertices) displayed as an actor in green. Matter propagates on a fixed background spacetime g_0 according to well-defined rules, particles know exactly what timelike geodesics are etc. This fixed background spacetime g_0 is displayed as a firm stage in blue. This is the situation of a QFT on a *Background Spacetime*.

In contrast, in Fig. 4 the stage has evaporated, it has become itself an actor (notice the arbitrarily high valent graviton (g) vertices) displayed in blue as well. Both matter and geometry are now dynamical entities and interact as displayed by the red vertex. There are no light cones any longer, rather the causal structure is a semiclassical concept only. This is the situation of a QFT on a *Differential Manifold* and this is precisely what QGR aims to rigorously define.

It is clear from these figures that the passage from a QFT on a background spacetime to a QFT on a differential manifold is a very radical one: It is like removing the chair on which you sit and trying to find a new, yet unknown, mechanism that keeps you from falling down. We should mention here that for many researchers in quantum gravity even that picture is not yet radical enough, some proposals require not only to get rid of the background metric g_0 but also of the differential manifold, allowing for topology change. This is also very desired in QGR but considered as a second step. In 3d QGR also this step could be completed and the final picture is completely combinatorial.

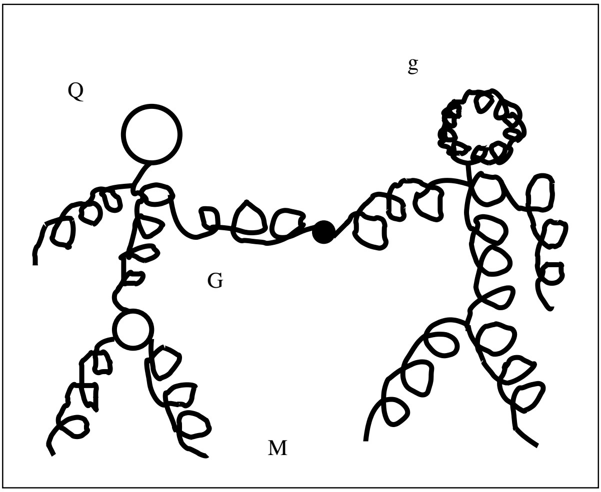


Fig. 4. QGR on Differential Manifold M: Actor = Matter + Geometry; Stage = Manifold M

Let us finish this section by stating once more what we mean by Quantum General Relativity (QGR).

Definition:
(Canonical) Quantum General Relativity (QGR) is an attempt to construct a mathematically rigorous, non-perturbative, background independent Quantum Field Theory of four-dimensional, Lorentzian General Relativity and all known matter in the continuum.
No additional, experimentally unverified structures are introduced. The fundamental principles of General Covariance and Quantum Theory are brought together and driven to their logical frontiers guided by mathematical consistency.
QGR is not a unified theory of all interactions in the standard sense since unification of gauge symmetry groups is not necessarily required in a non-perturbative approach. However, Geometry and Matter are unified in a non-standard sense by making them both transform covariantly under the Diffeomorphism Group at the quantum level.

1.2 Introduction: Classical Canonical Formulation of General Relativity

In this section we sketch the classical Hamiltonian formulation of general relativity in terms of Ashtekar's new variables. There are many ways to arrive at this new formulation and we will choose the one that is the most convenient one for our purposes.

The Hamiltonian formulation by definition requires some kind of split of the spacetime variables into time and spatial variables. This seems to contradict the whole idea of general covariance, however, quantum mechanics as presently formulated requires a notion of time because we interpret expectation values of operators as instantaneous measurement values averaged over a large number of measurements. In order to avoid this one has to "covariantize" the interpretation of quantum mechanics, in particular the measurement process, see e.g. [10] for a discussion. There are a number of proposals to make the canonical formulation more covariant, e.g.[4]: Multisymplectic Ansätze [13] in which there are multimomenta, one for each spacetime dimension, rather than just one for the time coordinate; Covariant phase space formulations [14] where one works on the space of solutions to the field equations rather than on the initial value instantaneous phase space; Peierl's bracket formulations [15] which covariantize the notion of the usual Poisson bracket; history bracket formulations [16], which grew out of the consistent history formulation of quantum mechanics [17], and which extends the usual spatial Poisson bracket to spacetime.

At the classical level all these formulations are equivalent. However, at the quantum level, one presently gets farthest within the the standard canonical formulation: The quantization of the multisymplectic approach is still in its beginning, see [18] for the most advanced results in this respect; The covariant phase space formulation is not only very implicit because one usually does not know the space of solutions to the classical field equations, but even if one manages to base a quantum theory on it, it will be too close to the classical theory since certainly the singularities of the classical theory are also built into the quantum theory; The Peierl's bracket also needs the explicit space of solutions to the classical field equations; Also the quantization of the history bracket formulation just has started, see [19] for first steps in that direction.

Given this present status of affairs, we will therefore proceed with the standard canonical quantization and see how far we get. Notice that there is no obvious problem with general covariance: For instance, standard Maxwell theory can be quantized canonically without any problem and one can show that the theory is Lorentz covariant although the spacetime split into space and time seems to break the Lorentz group down to the rotation group. *This is not at all the case!* It is just that Lorentz covariance is not manifest, one has to do some work in order to establish Lorentz covariance. Indeed, as we will see, at least at the classical level we will explicitly recover the four-dimensional diffeomorphism

[4] Path integrals [11] use the Lagrangean rather than the Hamiltonian and therefore seem to be better suited to a covariant formulation than the canonical one, however, usually the path integral is interpreted as some sort of propagator which makes use of instantaneous time Hilbert spaces again which therefore cannot be completely discarded with. At present, this connection with the canonical formulation is not very transparent, part of the reason being that the path integral is usually only defined in its Euclidean formulation, however the very notion of analytic continuation in time is not very meaningful in a theory where there is no distinguished choice of time, see however [12] for recent progress in this direction.

group in the formalism, although it is admittedly deeply hidden in the canonical formalism.

With these cautionary remarks out of the way, we will thus assume that the four dimensional spacetime manifold has the topology $\mathbb{R} \times \sigma$, where σ is a three dimensional manifold of arbitrary topology, in order to perform the $3+1$ split. This assumption about the topology of M may seem rather restrictive, however, it is not due to the following reasons: (1) According to a theorem due to Geroch any globally hyperbolic manifold (roughly those that admit a smooth metric with everywhere Lorentzian signature) is necessarily of that topology. Since Lorentzian metrics are what we are interested in, at least classically, the assumption about the topology of M is forced on us. (2) Any four manifold M has the topology of a countable disjoint union $\cup_\alpha I_\alpha \times \sigma_\alpha$ where either I_α are open intervals and σ_α is a three manifold or I_α is a one point set and σ_α is a two manifold (the latter are the intersections of the closures of the former). In this most generic situation we thus allow topology change between different three manifolds and it is even classically an open question how to make this compatible with the action principle. We take here a practical point of view and try to understand the quantum theory first for a single copy of the form $\mathbb{R} \times \sigma$ and later on worry how we glue the theories for different $\sigma' s$ together.

The ADM Formulation. In this nice situation the $3+1$ split is well known as the Arnowitt–Deser–Misner (ADM) formulation of general relativity, see e.g. [4] and we briefly sketch how this works. Since M is diffeomorphic to $\mathbb{R} \times \sigma$ we know that M foliates into hypersurfaces $\Sigma_t,\ t \in \mathbb{R}$ as in Fig. 5, where t labels the hypersurface and will play the role of our time coordinate. If we denote the four dimensional coordinates by $X^\mu,\ \mu = 0,1,2,3$ and the three dimensional coordinates by $x^a,\ a = 1,2,3$ then we know that there is a diffeomorphism $\varphi:\ \mathbb{R} \times \sigma \to M;\ (t,x) \mapsto X = \varphi(t,x)$ where $\Sigma_t = \varphi(t,\sigma)$. We stress that the four diffeomorphism φ is *completely arbitrary* until this point and thus *the*

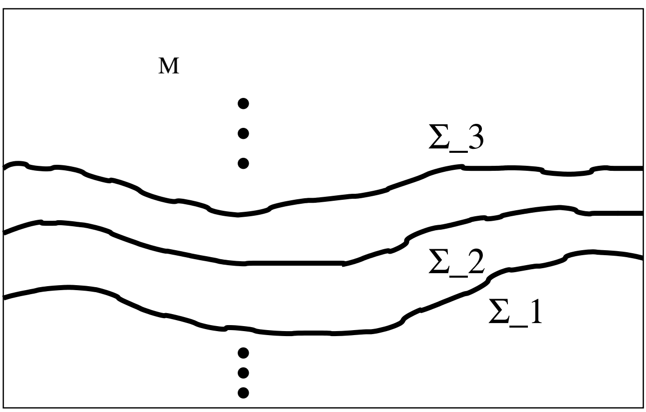


Fig. 5. Foliation of M

foliation of M is not at all fixed. In fact, when varying the diffeomorphism φ we obtain all possible foliations and the parametrization in terms of σ of each leaf Σ_t of the foliation can vary smoothly with φ. Consider the tangential vector fields to Σ_t given by

$$S_a(X) := (\partial_a)_{\varphi(t,x)=X} = (\varphi^\mu_{,a}(t,x))_{\varphi(t,x)=X}\ \partial_\mu \tag{3}$$

Denoting the four metric by $g_{\mu\nu}$ we define a normal vector field $n^\mu(X)$ by $g_{\mu\nu}n^\mu S^\nu_a = 0,\ g_{\mu\nu}n^\mu n^\nu = -1$. Thus, while the tangential vector fields depend only on the foliation, the normal vector field depends also on the metric. Let us introduce the foliation vector field

$$T(X) := (\partial_t)_{\varphi(t,x)=X} = (\varphi^\mu_{,t}(t,x))_{\varphi(t,x)=X}\ \partial_\mu \tag{4}$$

and let us decompose it into the basis n, S_a. This results in

$$T = Nn + U^a S_a \tag{5}$$

where N is called the *lapse function* while $U^a S_a$ is called the *shift vector* field. The arbitrariness of the foliation is expressed in the arbitrariness of the fields N, U^a. We can now introduce two symmetric spacetime tensor fields (∇ is the unique, torsion free covariant differential compatible with $g_{\mu\nu}$)

$$q_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu,\ \ K_{\mu\nu} = q_{\mu\rho}q_{\nu\sigma}\nabla^\rho n^\sigma \tag{6}$$

called the intrinsic metric and the extrinsic curvature respectively which are spatial, that is, their contraction with n vanishes. Thus, their full information is contained in their components with respect to the spatial fields S_a, e.g. $q_{ab}(t,x) = [q_{\mu\nu}S^\mu_a S^\nu_b](X(t,x))$. In particular,

$$K_{ab}(t,x) = \frac{1}{2N}[\dot{q}_{ab} - \mathcal{L}_U q_{ab}] \tag{7}$$

contains information about the *velocity* of q_{ab}. Here $\mathcal{L}$ the Lie derivative. The metric $g_{\mu\nu}$ is completely specified in terms of q_{ab}, N, U^a as one easily sees by expressing the line element $ds^2 = g_{\mu\nu}dX^\mu\, dX^\nu$ in terms of dt, dx^a.

Exercise 1.
Recall the definition of the Lie derivative and verify that $K_{\mu\nu}$ is indeed symmetric and that formula (7) holds.
Hint: A hypersurface Σ_t can be defined by the solution of an equation of the form $\tau(X) = t$. Conclude that $n_\mu \propto \nabla_\mu \tau$ and use torsion-freeness of ∇.

The Legendre transformation of the Einstein-Hilbert action

$$S = \frac{1}{\kappa}\int_M d^4X \sqrt{|\det(g)|}R^{(4)} \tag{8}$$

with q_{ab}, N, U^a considered as configuration coordinates in an infinite dimensional phase space is standard and we will not repeat the analysis here, which uses the so - called Gauss - Codazzi equations.

Here we are considering for simplicity only the case that σ is compact without boundary, otherwise (8) would contain boundary terms. The end result is

$$S = \frac{1}{\kappa}\int_{\mathbb{R}} dt \int_\sigma d^3x\{\dot{q}_{ab}P^{ab} + \dot{N}P + \dot{N}^a P_a - [\lambda P + \lambda^a P_a + U^a V_a + NC]\}\,, \quad (9)$$

where

$$P^{ab} = \frac{\kappa\delta S}{\delta q_{ab}} = \sqrt{\det(q)}[q^{ac}q^{bd} - q^{ab}q^{cd}]K_{cd} \quad (10)$$

and P, P_a are the momenta conjugate to q_{ab}, N, U^a respectively. Thus, we have for instance the equal time Poisson brackets

$$\begin{aligned}\{P^{ab}(t,x), P^{cd}(t,y)\} &= \{q_{ab}(t,x), q_{cd}(t,y)\} = 0\\ \{P^{ab}(t,x), q_{cd}(t,y)\} &= \kappa\delta^a_{(c}\delta^b_{d)}\delta(x,y)\end{aligned} \quad (11)$$

where $(.)_{(ab)} := [(.)_{ab} + (.)_{ba}]/2$ denotes symmetrization. The functions C, V_a which depend only on q_{ab}, P^{ab} are called the *Hamiltonian and Spatial Diffeomorphism constraint* respectively for reasons that will become obvious in a moment. Their explicit form is given by

$$\begin{aligned}V_a &= -2q_{ac}D_bP^{bc}\\ C &= \frac{1}{\sqrt{\det(q)}}[q_{ac}q_{bd} - \frac{1}{2}q_{ab}q_{cd}]P^{ab}P^{cd} - \sqrt{\det(q)}R\end{aligned} \quad (12)$$

where D is the unique, torsion-free covariant differential compatible with q_{ab} and R is the curvature scalar associated with q_{ab}.

The reason for the occurrence of the Lagrange multipliers λ, λ^a is that the Lagrangean (8) is singular, that is, one cannot solve all the velocities in terms of momenta and therefore one must use Dirac's procedure [20] for the Legendre transform of singular Lagrangeans. In this case the singularity structure is such that the momenta conjugate to N, U^a vanish identically, whence the Lagrange multipliers which when varied give the equations of motion $P = P_a = 0$. The equations of motion with respect to the Hamiltonian (i.e. $\dot{F} := \{H, F\}$ for any functional F of the canonical coordinates)

$$H = \int d^3x[\lambda P + \lambda^a P_a + U^a V_a + NC] \quad (13)$$

for N, U^a reveal that N, U^a are themselves Lagrange multipliers, i.e. completely unspecified functions (proportional to λ, λ^a) while the equations of motion for P, P_a give $\dot{P} = -C,\ \dot{P}_a = -V_a$. Since P, P_a are supposed to vanish, this requires $C = V_a = 0$ as well. Thus we see that the Hamiltonian *is constrained to vanish in GR!* We will see that this is a direct consequence of the four dimensional diffeomorphism invariance of the theory.

Now the equations of motion for q_{ab}, P^{ab} imply the so-called *Dirac (or hypersurface deformation) algebra*

$$\begin{aligned}\{V(U), V(U')\} &= \kappa V(\mathcal{L}_U U')\\ \{V(U), C(N)\} &= \kappa C(\mathcal{L}_U N)\\ \{C(N), C(N')\} &= \kappa V(q^{-1}(NdN' - N'dN))\,,\end{aligned} \quad (14)$$

where e.g. $C(N)=\int d^3xNC$. These equations tell us that the condition $H=V_a=0$ is preserved under evolution, in other words, the evolution is consistent! This is a non-trivial result. One says, the Hamiltonian and vector constraint form a *first class constraint algebra*. This algebra is much more complicated than the more familiar Kac-Moody algebras due to the fact that it is not an (infinite) dimensional Lie algebra in the true sense of the word because the "structure constants" on the right hand side of the last line in (14) are not really constants, they depend on the phase space. Such algebras are open in the the terminology of BRST [21] and about their representation theory only very little is known.

EXERCISE 2.
Derive (14) from (11).
Hint: Show first that the Poisson bracket between local functions which contain spatial derivatives is simply the spatial derivatives applied to the Poisson bracket. Since the Poisson bracket of local functions is distributional recall that derivatives of distributions are defined through an integration by parts.

Since the variables P, P_a drop out completely from the analysis and N, U^a are Lagrange multipliers, we may replace (9) by

$$S=\frac{1}{\kappa}\int_{\mathbb{R}} dt \int_\sigma d^3x\{\dot{q}_{ab}P^{ab}-[U^aV_a+NH]\} \tag{15}$$

with the understanding that N, U^a are now completely arbitrary functions which parameterize the freedom in choosing the foliation. Since the Hamiltonian of GR depends on the completely unspecified functions N, U^a, the motions that it generates in the phase space $\mathcal{M}$ coordinatized by (P^{ab}, q_{ab}) subject to the Poisson brackets (11) are to be considered as pure gauge transformations. The infinitesimal flow (or motion) of the canonical coordinates generated by the corresponding Hamiltonian vector fields on $\mathcal{M}$ has the following form for an arbitrary tensor t_{ab} built from q_{ab}, P^{ab}

$$\begin{aligned}\{V(U),t_{ab}\}_{EOM}&=\kappa(\mathcal{L}_U t_{ab})\\ \{C(N),t_{ab}\}_{EOM}&=\kappa(\mathcal{L}_{Nn} t_{ab})\,,\end{aligned} \tag{16}$$

where the subscript EOM means that these relations hold for generic functions on $\mathcal{M}$ only when the vacuum equations of motion (EOM) $R^{(4)}_{\mu\nu}-R^{(4)}g_{\mu\nu}/2=0$ hold. Equation (16) reveals that Diff(M) is implemented also in the canonical formalism, however, in a rather non-trivial way: The gauge motions generated by the constraints can be interpreted as four-dimensional diffeomorphisms *only when the EOM hold*. This was to be expected because a diffeomorphism orthogonal to the hypersurface means evolution in the time parameter, what is surprising though is that this evolution is considered as a gauge transformation in GR. Off the solutions, the constraints generate different motions, in other words, the set of gauge symmetries is not Diff(M) everywhere in the phase space. This is not unexpected: The action (8)narray is obviously Diff(M) invariant, but so would be any action that is an integral over a four-dimensional scalar density of weight one

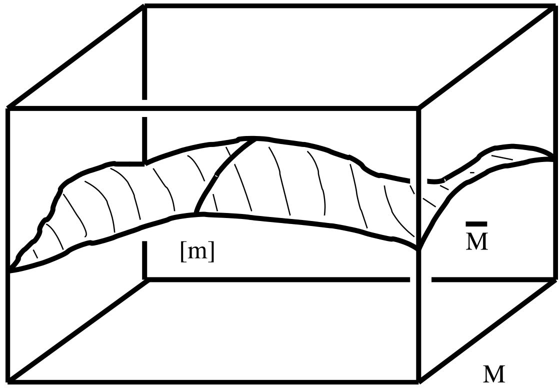


Fig. 6. Constraint submanifold $\overline{\mathcal{M}}$ and gauge orbit $[m]$ of $m \in \overline{\mathcal{M}}$ in $\mathcal{M}$

formed from polynomials in the curvature tensor and its covariant derivatives. This symmetry is completely insensitive to the specific Lagrangean in question, it is kinematical. The dynamics generated by a specific Lagrangean must depend on that Lagrangean, otherwise all Lagrangeans underlying four dimensionally diffeomorphism invariant actions would equal each other up to a diffeomorphism which is certainly not the case (consider for instance higher derivative theories). In particular, that dynamics is, a priori, completely independent of $\mathrm{Diff}(M)$. As a consequence, Dirac observables, that is, functions on $\mathcal{M}$ which are gauge invariant (have vanishing Poisson brackets with the constraints), *are not simply functionals of the four metric invariant under four diffeomorphisms* because they must depend on the Lagrangean. The set of these dynamics dependent gauge transformations does not obviously form a group as has been investigated by Bergmann and Komar [22]. The geometrical origin of the hypersurface deformation algebra has been investigated in [23]. Torre and Anderson have shown that for compact σ there are no Dirac observables which depend on only a finite number of spatial derivatives of the canonical coordinates [24] which means that Dirac observables will be highly non-trivial to construct.

Let us summarize the gauge theory of GR in Fig. 6: The constraints $C = V_a = 0$ define a constraint submanifold $\overline{\mathcal{M}}$ within the full phase space $\mathcal{M}$. The gauge motions are defined on all of $\mathcal{M}$ but they have the feature that they leave the constraint submanifold invariant, and thus the orbit of a point m in the submanifold under gauge transformations will be a curve or gauge orbit $[m]$ entirely within it. The set of these curves defines the so-called reduced phase space and Dirac observables restricted to $\overline{\mathcal{M}}$ depend only on these orbits. Notice that as far as the counting is concerned we have twelve phase space coordinates q_{ab}, P^{ab} to begin with. The four constraints C, V_a can be solved to eliminate four of those and there are still identifications under four independent sets of motions among the remaining eight variables leaving us with only four Dirac observables. The corresponding so-called reduced phase space has therefore precisely the two configuration degrees of freedom of general relativity.

Gauge Theory Formulation. We can now easily introduce the shift from the ADM variables q_{ab}, P^{ab} to the connection variables introduced first by Ashtekar [25] and later somewhat generalized by Immirzi [26] and Barbero [27]. We introduce $su(2)$ indices $i, j, k, .. = 1, 2, 3$ and co-triad variables e_a^j with inverse e_j^a whose relation with q_{ab} is given by

$$q_{ab} := \delta_{jk} e_a^j e_b^k \tag{17}$$

Defining the spin connection Γ_a^j through the equation

$$\partial_a e_b^j - \Gamma_{ab}^c e_c^j + \epsilon_{jkl} \Gamma_a^k e_b^l = 0 \,, \tag{18}$$

where Γ_{ab}^c are the Christoffel symbols associated with q_{ab} we now define

$$A_a^j = \Gamma_a^j + \beta K_{ab} e_j^b, \quad E_j^a = \sqrt{\det(q)} e_j^a / \beta \,, \tag{19}$$

where $\beta \in \mathbb{C} - \{0\}$ is called the *Immirzi* parameter. In this article we only consider real valued and positive β. Finally we introduce the $SU(2)$ *Gauss constraint*

$$G_j := \partial_a E_j^a + \epsilon_{jkl} A_a^k E_l^b \tag{20}$$

with ϵ_{jkl} the structure constants of $su(2)$ which we would encounter in the canonical formulation of any $SU(2)$ gauge theory. As one can check, modulo $G_j = 0$ one can then write C, V_a in terms of A, E as follows

$$\begin{aligned} V_a &= F_{ab}^j E_j^b, \\ C &= \frac{F_{ab}^j \epsilon_{jkl} E_j^a E_l^b}{\sqrt{|\det(E)|}} \quad + \quad \text{More} \;, \end{aligned} \tag{21}$$

where $F = 2(dA + A \wedge A)$ is the curvature of A and "More" is an additional term which is more complicated but can be treated by similar methods as the one displayed.

We then have the following theorem [25].

THEOREM 2.
Consider the phase space $\mathcal{M}$ coordinatized by (A_a^j, E_j^b) with Poisson brackets

$$\{E_j^a(x), E_k^b(y)\} = \{A_a^j(x), A_b^k(y)\} = 0, \;\; \{E_j^a(x), A_b^k(y)\} = \kappa \delta_b^a \delta_j^k \delta(x,y) \tag{22}$$

and constraints G_j, C, V_a. Then, solving only the constraint $G_j = 0$ and determining the Dirac observables with respect to it leads us back to the ADM phase space with constraints C, V_a.

The proof of the theorem is non-trivial and tedious and can be found in the notation used here in [1]. Alternatively one can find directions for a proof in the subsequent exercise. In particular, this works only because the Gauss constraint is in involution with itself and the other constraints, specifically $\{G, G\} \propto G, \; \{G, V\} = \{G, H\} = 0$.

EXERCISE 3.

i) Prove theorem 2.
Hint: Express q_{ab}, P^{ab} in terms of A_a^j, E_j^a by using (17), (18), (19), and (20) and check that the Poisson brackets, with respect to (22), among the solutions $q_{ab} = s_{ab}[A, E], P^{ab} = S^{ab}[A, E]$ equal precisely (11) modulo terms proportional to G_j.

ii) Define $G(\Lambda) := \int_\sigma d^3x \Lambda^j G_j$, $D(U) := \int_\sigma U^a[V_a - A_a^j G_j]$, and $[\Lambda, \Lambda']_j = \epsilon_{jkl}\Lambda^k(\Lambda')^l$. Verify the following Poisson brackets

$$\begin{aligned}
\{G(\Lambda), G(\Lambda')\} &= \kappa G([\Lambda, \Lambda']) \\
\{G(\Lambda), V(U)\} &= 0 \\
\{D(U), D(U')\} &= \kappa D([U, U'])
\end{aligned} \tag{23}$$

and conclude that the Hamiltonian vector fields of $G(\Lambda)$ and $D(U)$ respectively generate $SU(2)$ gauge transformations and spatial diffeomorphisms of σ respectively.
Hint: Show first that

$$\begin{aligned}
\{G(\Lambda/\kappa), A_a^j(x)\} &= -\Lambda^j_{,a} + \epsilon_{jkl}\Lambda^k A_a^l \\
\{D(U/\kappa), A_a^j(x)\} &= U^b A^j_{a,b} + U^b_{,a} A_b^j
\end{aligned} \tag{24}$$

to conclude that A transforms as a connection under infinitesimal gauge transformations and as a one-form under infinitesimal diffeomorphisms. Consider then $g_t(x) := \exp(t\Lambda^j\tau_j/(2\kappa))$ and $\varphi_t(x) := c_{U,x}(t)$ where $t \mapsto c_{U,x}(t)$ is the unique integral curve of U through x, that is, $\dot{c}_{U,x}(t) = U(c_{U,x}(t))$, $c_{U,x}(0) = x$. Recall that the usual transformation behaviour of connections and one-forms under finite gauge transformations and diffeomorphisms respectively is given by (e.g. [28])

$$\begin{aligned}
A^g &= -dgg^{-1} + \mathrm{Ad}_g(A) \\
A^\varphi &= \varphi^* A\,,
\end{aligned} \tag{25}$$

where $A = A_a^j dx^a \tau_j/2$, $\mathrm{Ad}_g(.) = g(.)g^{-1}$ denotes the adjoint representation of $SU(2)$ on $su(2)$ and φ^* denotes the pull-back map of p-forms and $i\tau_j$ are the Pauli matrices so that $\tau_j\tau_k = -\delta_{jk}1_2 + \epsilon_{jkl}\tau_l$. Verify then that (24) is the derivative at $t = 0$ of (25) with $g := g_t$, $\varphi := \varphi_t$. Similarly, derive that E transforms as an $su(2)$-valued vector field of density weight one. (Recall that a tensor field t of some type is said to be of density weight $r \in \mathbb{R}$ if $t\sqrt{|\det(s)|}^{-r}$ is an ordinary tensor field of the same type where s_{ab} is any non-degenerate symmetric tensor field).

From the point of view of the classical theory we have made things more complicated: Instead of twelve variables q, P we now have eighteen A, E. However, the additional six phase-space dimensions (per spacepoint) are removed by the first class Gauss constraint which shows that working on our gauge theory phase space is equivalent to working on the ADM phase space. The virtue of this extended phase space is that *canonical GR can be formulated in the language of a canonical gauge theory* where A plays the role of an $SU(2)$ connection with canonically conjugate electric field E. Besides the remark that this fact could be the starting point for a possible gauge group unification of all four forces we now have access to a huge arsenal of techniques that have been developed for the canonical quantization of gauge theories. It is precisely this fact that has enabled steady progress in this field in the last fifteen years while one was stuck with the ADM formulation for almost thirty years.

1.3 Canonical Quantization Programme for Theories with Constraints

Refined Algebraic Quantization (RAQ). As we have seen, GR can be formulated as a constrained Hamiltonian system with first class constraints. The quantization of such systems has been considered first by Dirac [20] and was later refined by a number of authors. It is now known under the name refined algebraic quantization (RAQ). We will briefly sketch the main ideas following [29].

i) *Phase Space and Constraints*
The starting point is a phase space $(\mathcal{M}, \{.,.\})$ together with a set of first class constraints C_I and possibly a Hamiltonian H.

ii) *Choice of Polarization*
In order to quantize the phase space we must choose a polarization, that is, a Lagrangean submanifold $\mathcal{C}$ of $\mathcal{M}$ which is called configuration space. The coordinates of $\mathcal{C}$ have vanishing Poisson brackets among themselves. If $\mathcal{M}$ is a cotangent bundle, that is, $\mathcal{M} = T^*\mathcal{Q}$ then it is natural to choose $\mathcal{Q} = \mathcal{C}$ and we will assume this to be the case in what follows. For more general cases, e.g. compact phases spaces one needs ideas from geometrical quantization, see e.g. [30]. The idea is that (generalized, see below) points of $\mathcal{C}$ serve as arguments of the vectors of the Hilbert space to be constructed.

iii) *Preferred Kinematical Poisson Subalgebra*
Consider the space $C^\infty(\mathcal{C})$ of smooth functions on $\mathcal{C}$ and the space $V^\infty(\mathcal{C})$ of smooth vector fields on C. The vertical polarization of $\mathcal{M}$, that is, the space of fibre coordinates called momentum space, generates preferred elements of $V^\infty(\mathcal{C})$ through $(v_p[f])(q) := (\{p, f\})(q)$ where we have denoted configuration and momentum coordinates by q, p respectively and $v[f]$ denotes the action of a vector field on a function. The pair $C^\infty(\mathcal{C}) \times \mathcal{V}^\infty(\mathcal{C})$ forms a Lie algebra defined by $[(f, v), (f', v')] = (v[f'] - v'[f], [v, v'])$ of which the algebra $\mathcal{B}$ generated by elements of the form (f, v_p) forms a subalgebra. We assume that $\mathcal{B}$ is closed under complex conjugation which becomes its *-operation (involution).

iv) *Representation Theory of the Corresponding Abstract *-Algebra*
We are looking for all irreducible *-representations $\pi : \mathcal{B} \to \mathcal{L}(\mathcal{H}_{\text{kin}})$ of $\mathcal{B}$ as linear operators on a kinematical Hilbert space $\mathcal{H}_{\text{kin}}$ such that the *-relations becomes the operator adjoint and such that the canonical commutation relations are implemented, that is, for all $a, b \in \mathcal{B}$

$$\begin{aligned} \pi(a)^\dagger &= \pi(a^*) \\ [\pi(a), \pi(b)] &= i\hbar\pi([a, b]) \,. \end{aligned} \tag{26}$$

Strictly speaking, (26) is to be supplemented by the domains on which the operators are defined. In order to avoid this one will work with the subalgebra of $C^\infty(\mathcal{C})$ formed by bounded functions, say of compact support and one will deal with exponentiated vector fields in order to obtain bounded operators. Irreducibility is a physically meaningful requirement because we are not interested in Hilbert spaces with superselection sectors and the reason for why we do not require the full Poisson algebra to be faithfully represented is that this is almost always impossible in irreducible representations as stated in the famous Groenewald–van Hove theorem (compare Giulini's contribution to this volume). The Hilbert space that one gets can usually be described in the form $L_2(\overline{\mathcal{C}}, d\mu)$ where $\overline{\mathcal{C}}$ is a distributional extension of $\mathcal{C}$ and μ is a probability measure thereon. A well-known example is the case of free scalar fields on Minkowski space where $\mathcal{C}$ is some space of smooth scalar fields on $\mathbb{R}^3$ vanishing at spatial infinity while $\overline{\mathcal{C}}$ is the space of tempered distributions on $\mathbb{R}^3$ and μ is a normalized Gaussian measure on $\overline{\mathcal{C}}$.

v) *Selection of Suitable Kinematical Representations*
Certainly we want a representation which supports also the constraints and the Hamiltonian as operators which usually will limit the number of available representations to a small number, if possible at all. The constraints usually are not in $\mathcal{B}$ unless linear in momentum and the expressions $\hat{C}_I := \pi(C_I), \hat{H} = \pi(H)$ will involve factor ordering ambiguities as well as regularization and renormalization processes in the case of field theories. In the generic case, $\hat{C}_I, \hat{H}$ will not be bounded and $\hat{C}_I$ will not be symmetric. We will require that $\hat{H}$ is symmetric and that the constraints are at least closable, that is, they are densely defined together with their adjoints. It is then usually not too difficult to find a dense domain $\mathcal{D}_{\text{kin}} \subset \mathcal{H}_{\text{kin}}$ on which all these operators and their adjoints are defined and which they leave invariant. Typically $\mathcal{D}_{\text{kin}}$ will be a space of smooth functions of rapid decrease so that arbitrary derivatives and polynomials of the configuration variables are defined on them and such spaces naturally come with their own topology which is finer than the subspace topology induced from $\mathcal{H}_{\text{kin}}$ whence we have a topological inclusion $\mathcal{D}_{\text{kin}} \hookrightarrow \mathcal{H}_{\text{kin}}$.

vi) *Imposition of the Constraints*
The two step process in the classical theory of solving the constraints $C_I = 0$ and looking for the gauge orbits is replaced by a one step process in the quantum theory, namely looking for solutions l of the equations $\hat{C}_I l = 0$. This is because it obviously solves the constraint at the quantum level (in the corresponding representation on the solution space the constraints are replaced by the zero operator) and it simultaneously looks for states that are gauge invariant because $\hat{C}_I$ is the quantum generator of gauge transformations.

Now, unless the point $\{0\}$ is in the common point spectrum of all the $\hat{C}_I$, solutions l to the equations $\hat{C}_I l = 0 \; \forall \; I$ do not lie in $\mathcal{H}_{\text{kin}}$, rather they are distributions. Here one has several options, one could look for solutions in the space $\mathcal{D}'_{\text{kin}}$ of continuous linear functionals on $\mathcal{D}_{\text{kin}}$ (topological dual) or in the space $\mathcal{D}^*_{\text{kin}}$ of linear functionals on $\mathcal{D}_{\text{kin}}$ with the topology of pointwise convergence (algebraic dual). Since certainly $\mathcal{H}_{\text{kin}} \subset \mathcal{D}'_{\text{kin}} \subset \mathcal{D}^*_{\text{kin}}$ let us choose the latter option for the sake of more generality. The topology on $\mathcal{H}_{\text{kin}}$ is finer than the subspace topology induced from $\mathcal{D}^*_{\text{kin}}$ so that we obtain a *Gel'fand triple* or *Rigged Hilbert Space*

$$\mathcal{D}_{\text{kin}} \hookrightarrow \mathcal{H}_{\text{kin}} \hookrightarrow \mathcal{D}^*_{\text{kin}} \,. \tag{27}$$

This a slight abuse of terminology since the name is usually reserved for the case that $\mathcal{D}_{\rm kin}$ carries a nuclear topology (generated by a countable family of seminorms separating the points) and that $\mathcal{D}^*_{\rm kin}$ is its topological dual.
We are now looking for a subspace $\mathcal{D}^*_{\rm phys} \subset \mathcal{D}^*_{\rm kin}$ such that for its elements l holds

$$[\hat{C}'_I l](f) := l(\hat{C}_I^\dagger f) = 0 \;\; \forall \, f \in \mathcal{D}_{\rm kin}, \; \forall I \, . \tag{28}$$

The prime on the left hand side of this equation defines a dual, anti-linear representation of the constraints on $\mathcal{D}^*_{\rm kin}$. The reason for the adjoint on the right hand side of this equation is that if l would be an element of $\mathcal{H}_{\rm kin}$ then (28) would be replaced by

$$[\hat{C}'_I l](f) :=< \hat{C}_I l, f >_{\rm kin} =< l, \hat{C}_I^\dagger f >_{\rm kin} =: l(\hat{C}_I^\dagger f) \;\; \forall \, f \in \mathcal{D}_{\rm kin}, \; \forall I \, , \tag{29}$$

where $< .,. >_{\rm kin}$ denotes the kinematical inner product, so that (28) is the natural extension of (29) from $\mathcal{H}_{\rm kin}$ to $\mathcal{D}^*_{\rm kin}$.

vii) *Anomalies*
Since we have a first class constraint algebra, we know that classically $\{C_I, C_J\} = f_{IJ}\;^K C_K$ for some *structure functions* $f_{IJ}\;^K$ which depend in general on the phase space point $m \in \mathcal{M}$. The translation of this equation into quantum theory is then plagued with ordering ambiguities, because the structure functions turn into operators as well. It may therefore happen that, e.g.

$$[\hat{C}_I, \hat{C}_J] = i\hbar \hat{C}_K \hat{f}_{IJ}\;^K = i\hbar\{[\hat{C}_K, \hat{f}_{IJ}\;^K] + \hat{f}_{IJ}\;^K \hat{C}_K\} \tag{30}$$

and it follows that any $l \in \mathcal{D}^*_{\rm phys}$ also solves the equation $([\hat{C}_K, \hat{f}_{IJ}\;^K])' l = 0$ for all I, J. If that commutator is not itself a constraint again, then it follows that l solves more than only the equations $\hat{C}'_I l = 0$ and thus the quantum theory has less physical degrees of freedom than the classical theory. This situation, called an *anomaly*, must be avoided by all means.

viii) *Dirac Observables and Physical Inner Product*
Since generically $\mathcal{H}_{\rm kin} \cap \mathcal{D}^*_{\rm phys} = \emptyset$, the space $\mathcal{D}^*_{\rm phys}$ cannot be equipped with the scalar product $< .,. >_{\rm kin}$. It is here where Dirac observables come into play. A *strong* Dirac observable is an operator $\hat{O}$ on $\mathcal{H}_{\rm kin}$ which is, together with its adjoint, densely defined on $\mathcal{D}_{\rm kin}$ and which commutes with all constraints, that is, $[\hat{O}, \hat{C}_I] = 0$ for all I. We require that $\hat{O}$ is the quantization of a real valued function O on the phase space and the condition just stated is the quantum version of the classical gauge invariance condition $\{O, C_I\} = 0$ for all I. A *weak* Dirac observable is the quantum version of the more general condition $\{O, C_I\}_{|C_J} = 0 \forall J = 0 \; \forall I$ and simply means that the space of solutions is left invariant by the natural dual action of the operator $\hat{O}' \mathcal{D}^*_{\rm phys} \subset \mathcal{D}^*_{\rm phys}$ (compare Giulini's contribution to this volume).
A physical inner product on a subset $\mathcal{H}_{\rm phys} \subset \mathcal{D}^*_{\rm phys}$ is a positive definite sesquilinear form $< .,. >_{\rm phys}$ with respect to which the $\hat{O}'$ become self-adjoint operators, that is, $\hat{O}' = (\hat{O}')^\star$ where the adjoint on $\mathcal{H}_{\rm phys}$ is denoted

by $\star$. Notice that $[\hat{O}_1', \hat{O}_2'] = ([\hat{O}_1, \hat{O}_2])'$ so that commutation relations on $\mathcal{H}_{\rm kin}$ are automatically transferred to $\mathcal{H}_{\rm phys}$ which then carries a proper *-representation of the physical observables. The observables themselves will only be defined on a dense domain $\mathcal{D}_{\rm phys} \subset \mathcal{H}_{\rm phys}$ and we get a second Gel'fand triple

$$\mathcal{D}_{\rm phys} \hookrightarrow \mathcal{H}_{\rm phys} \hookrightarrow \mathcal{D}^*_{\rm phys}. \tag{31}$$

In fortunate cases, for instance when the $\hat{C}_I$ are mutually commuting self-adjoint operators on $\mathcal{H}_{\rm kin}$, all we have said is just a fancy way of stating the fact that $\mathcal{H}_{\rm kin}$ has a direct integral decomposition

$$\mathcal{H}_{\rm kin} = \int_S^{\oplus} d\nu(\lambda)\mathcal{H}_\lambda \tag{32}$$

over the spectrum S of the constraint algebra with a measure ν and eigenspaces $\mathcal{H}_\lambda$ which are left invariant by the strong observables and therefore $\mathcal{H}_{\rm phys} = \mathcal{H}_0$. In the more general cases that are of concern to us, more work is required.

ix) *Classical Limit*
It is by no means granted that the representation $\mathcal{H}_{\rm phys}$ that one finally arrived at carries semiclassical states, that is states $\psi_{[m]}$ labelled by gauge equivalence classes $[m]$ of points $m \in \mathcal{M}$ with respect to which the Dirac observables have the correct expectation values and with respect to which their relative fluctuations are small, that is, roughly speaking

$$\left|\frac{<\psi_{[m]}, \hat{O}'\psi_{[m]}>_{\rm phys}}{O(m)} - 1\right| \ll 1 \text{ and } \left|\frac{<\psi_{[m]}, (\hat{O}')^2\psi_{[m]}>_{\rm phys}}{(<\psi_{[m]}, \hat{O}'\psi_{[m]}>_{\rm phys})^2} - 1\right| \ll 1. \tag{33}$$

Only when such a phase exists are we sure that we have not constructed some completely spurious sector of the quantum theory which does not admit the correct classical limit.

Selected Examples with First Class Constraints. In the case that a theory has only first class constraints, Dirac's algorithm [20] boils down to the following four steps:

1. Define the momentum p_a conjugate to the configuration variable q^a by (Legendre transform)

$$p_a := \partial S / \partial \dot{q}^a\,, \tag{34}$$

   where S is the action.
2. Equation (34) defines p_a as a function of $q^a, \dot{q}^a$ and if it is not invertible to define the $\dot{q}^a$ as a function of q^a, p_a we get a collection of so-called *primary constraints* C_I, that is, identities among the q^a, p_a. In this situation one says that S or the Lagrangean is singular.

3. Using that q^a, p_a have canonical Poisson brackets, compute all possible Poisson brackets $C_{IJ} := \{C_I, C_J\}$. If some $C_{I_0 J_0}$ is not zero when all C_K vanish, then add this $C_{I_0 J_0}$, called a *secondary constraint*, to the set of primary constraints.
4. Iterate 3) until the C_I are in involution, that is, no new secondary constraints appear.

In this report we will only deal with theories which have no second class constraints, so this algorithm is all we need.

Exercise 4.
Perform the quantization programme for a couple of simple systems in order to get a feeling for the formalism:

1. *Momentum Constraint*
$\mathcal{M} = T^*\mathbb{R}^2$ with standard Poisson brackets among q^a, p_a; $a = 1, 2$ and constraint $C := p_1$. Choose $\mathcal{H}_{\rm kin} = L_2(\mathbb{R}^2, d^2x)$, $\mathcal{D}_{\rm kin} = \mathcal{S}(\mathbb{R}^2)$, $\mathcal{D}^*_{\rm kin} = \mathcal{S}'(\mathbb{R}^2)$ (spaces of functions of rapid decrease and tempered distributions respectively).
Solution: Dirac observables are the conjugate pair q^2, p_2, $\mathcal{H}_{\rm phys} = L_2(\mathbb{R}, dx_2)$.
Hint: Work in the momentum representation and conclude that the general solution is of the form $l_f(p_1, p_2) = \delta(p_1) f(p_2)$ for $f \in \mathcal{S}'(\mathbb{R})$.
2. *Angular Momentum Constraint*
$\mathcal{M} = T^*\mathbb{R}^3$ with standard Poisson brackets among q^a, p_a; $a = 1, 2, 3$ and constraints $C_a := \epsilon_{abc} x^b p_c$. Check the first class property and choose the kinematical spaces as above with $\mathbb{R}^2$ replaced by $\mathbb{R}^3$.
Solution: Dirac observables are the conjugate pair $r := \sqrt{\delta_{ab} q^a q^b} \geq 0$, $p_r = \delta_{ab} q^a p_b / r$, the physical phase space is $T^*\mathbb{R}_+$ and $\mathcal{H}_{\rm phys} = L_2(\mathbb{R}_+, r^2 dr)$ where $\hat{r}$ is a multiplication operator and $\hat{p}_r = i\hbar \frac{1}{r} \frac{d}{dr} r$ with dense domain of symmetry given by the square integrable functions f such that f is regular at $r = 0$.
Hint: Introduce polar coordinates and decompose kinematical wave functions into spherical harmonics. Conclude that the physical Hilbert space this time is just the restriction of the kinematical Hilbert space to the zero angular momentum subspace, that is, $\mathcal{H}_{\rm phys} \subset \mathcal{H}_{\rm kin}$. The reason is of course that the spectrum of the $\hat{C}_a$ is pure point (discrete).
3. *Relativistic Particle*
Consider the Lagrangean $L = -m\sqrt{-\eta_{\mu\nu}\dot{q}^\mu \dot{q}^\nu}$ where m is a mass parameter, η is the Minkowski metric and $\mu = 0, 1, .., D$. Verify that the Lagrangean is singular, that is, the velocities $\dot{q}^\mu$ cannot be expressed in terms of the momenta $p_\mu = \partial L / \partial \dot{q}^\mu$ which gives rise to the *mass shell constraint* $C = m^2 + \eta^{\mu\nu} p_\mu p_\nu$. Verify that this happens because the corresponding action is invariant under Diff($\mathbb{R}$), that is, reparameterizations $t \mapsto \varphi(t)$, $\dot{\varphi}(t) > 0$. Perform the Dirac analysis for constraints and conclude that the system has no Hamiltonian, just the Hamiltonian constraint C which generates reparameterizations on the kinematical phase space $\mathcal{M} = T^*\mathbb{R}^{D+1}$ with standard Poisson brackets. Now choose kinematical spaces as in 1. with $\mathbb{R}^2$ replaced by $\mathbb{R}^{D+1}$.
Solution: Conjugate Dirac observables are
$$Q^a = q^a - \frac{q^0 p_a}{\sqrt{m^2 + \delta^{ab} p_a p_b}} \tag{35}$$

and

$$\mathcal{H}_{\rm phys} = L_2(\mathbb{R}^D, d^D p) \tag{36}$$

on which $\hat{q}^0 = 0$.
Hint: Work in the momentum representation and conclude that the general solution to the constraints is of the form $l_f = \delta(C) f(p_0, \boldsymbol{p})$. Now notice that the δ-distribution can be written as a sum of two δ-distribution corresponding to the positive and negative mass shell and choose f to have support in the former.
This example has features rather close to those of general relativity.

4. *Maxwell Theory*
Consider the action for free Maxwell-theory on Minkowski space and perform the Legendre transform. Conclude that there is a first class constraint $C = \partial_a E^a$ (Gauss constraint) with Lagrange multiplier A_0 and a Hamiltonian

$$H = \frac{1}{2} \int_{\mathbb{R}^3} d^3x (E^a E^b + B^a B^b) , \tag{37}$$

where $E^a = \dot{A}_a - \partial_a A_0$ is the electric field and $B^a = \epsilon^{abc} \partial_b A_c$ the magnetic one. Verify that the Gauss constraint generates $U(1)$ gauge transformations $A \mapsto A - df$ while E^a is gauge invariant. Choose $\mathcal{H}_{\rm kin}$ to be the standard Fock space for three massless, free scalar fields A_a and as $\mathcal{D}_{\rm kin}$, $\mathcal{D}^*_{\rm kin}$ the finite linear span of n-particle states and its algebraic dual respectively.
Solution: Conjugate Dirac observables are the transversal parts of A, E respectively, e.g. $E^a_\perp = E^a - \partial_a \frac{1}{\Delta} \partial_b E^b$ where Δ is the Laplacian on $\mathbb{R}^3$. The physical Hilbert space is the standard Fock space for two free, massless scalar fields corresponding to these transversal degrees of freedom.
Hint: Fourier transform the fields and compute the standard annihilation and creation operators $\hat{z}_a(k), \hat{z}^\dagger_a(k)$ with canonical commutation relations. Express the Gauss constraint operator in terms of them and conclude that the gauge invariant part satisfies $\hat{z}_a(k) k^a = 0$. Introduce $\hat{z}_I(k) = \hat{z}_a(k) e^a_I(k)$ where $\boldsymbol{e}_1(k), \boldsymbol{e}_2(k), \boldsymbol{e}_3(k) := \boldsymbol{k}/||k||$ form an oriented orthonormal basis. Conclude that physical states are states without longitudinal excitations and build the Fock space generated by the $\hat{z}^\dagger_1(k), \hat{z}^\dagger_2(k)$ from the kinematical vacuum state.

2 Mathematical and Physical Foundations of Quantum General Relativity

2.1 Mathematical Foundations

Polarization and Preferred Poisson Algebra $\boldsymbol{\mathcal{B}}$. The first two steps of the quantization programme were already completed in Sect. 1.2: The phase space $\mathcal{M}$ is coordinatized by canonically conjugate pairs (A^j_a, E^a_j) where A is an $SU(2)$ connection over σ while E is a $su(2)$-valued vector density of weight one over σ and the Poisson brackets were displayed in (22). Strictly speaking, since $\mathcal{M}$ is an infinite dimensional space, one must supply $\mathcal{M}$ with a manifold structure modelled on some Banach space but we will skip these functional analytic niceties here, see [1] for further information. Also we must specify the principal fibre bundle of which A is the pull-back by local sections of a globally defined connection, and we must specify the vector bundle associated to that principal bundle under the adjoint representation of which E is the pull-back by local sections. Again, in order not to dive too deeply into fibre bundle theoretic subtleties, we will assume that the principal fibre bundle is trivial so that A, E are actually globally defined. In fact, for the case of $G = SU(2)$ and $\dim(\sigma) = 3$ one can show that the fibre bundle is necessarily trivial but for the generalization to the generic case we again refer the reader to [1].

With these remarks out of the way we may begin by defining a polarization. The fact that GR has been casted into the language of a gauge theory suggests the choice $\mathcal{C} = \mathcal{A}$, the space of smooth $SU(2)$ connections over σ.

The next question then is how to choose the space $C^\infty(\mathcal{A})$. Since we are dealing with a field theory, it is not clear a priori what smooth or even differentiable means. In order to give precise meaning to this, one really has to equip $\mathcal{A}$ with a manifold structure modelled on a Banach space. This is because one usually says that a function $F : \mathcal{A} \to \mathbb{C}$ is differentiable at $A_0 \in \mathcal{A}$ provided that there exists a *bounded* linear functional $DF_{A_0} : T_{A_0}(\mathcal{A}) \to \mathbb{C}$ such that $F[A_0 + \delta A] - F[A_0] - DF_{A_0} \cdot \delta A$ vanishes "faster than linearly" for arbitrary tangent vectors $\delta A \in T_{A_0}(\mathcal{A})$ at A_0. (The proper way of saying this is using the natural Banach norm on $T(\mathcal{A})$.) Of course, in physicist's notation the differential $DF_{A_0} = (\delta F/\delta A)(A_0)$ is nothing else than the functional derivative. Using this definition it is clear that polynomials in $A_a^j(x)$ are not differentiable because their functional derivative is proportional to a δ-distribution as it is clear from (22). Thus we see that the smooth functions of A have to involve some kind of smearing of A with test functions, which is generic in field theories.

Now this smearing should be done in a judicious way. The function $F[A] := \int_\sigma d^3x F_j^a(x) A_a^j(x)$ for some smooth test function F_j^a of compact support is certainly smooth in the above sense, its functional derivative being equal to F_a^j (which is a bounded operator if F is e.g. an L_2 function on σ and the norm on the tangent spaces is an L_2 norm). However, this function does not transform nicely under $SU(2)$ gauge transformations which will make it very hard to construct $SU(2)$ invariant functions from them. Here it helps to look up how physicists have dealt with this problem in ordinary canonical quantum Yang-Mills gauge theories and they found the following, more or less unique solution [31]: Given a curve $c : [0,1] \to \sigma$ in σ and a point $A \in \mathcal{A}$ we define the holonomy or parallel transport $A(c) := h_{c,A}(1) \in SU(2)$ as the unique solution to the following ordinary differential equation for functions $h_{c,A} : [0,1] \to SU(2)$

$$\frac{dh_{c,A}(t)}{dt} = h_{c,A}(t) A_a^j(c(t)) \frac{\tau_j}{2} \dot{c}^a(t)\,, \qquad h_{c,A}(0) = 1_2\,. \tag{38}$$

Exercise 5.

Verify that (38) is equivalent with

$$A(c) = \mathcal{P} \cdot \exp\left(\int_c A\right) = 1_2 + \sum_{n=1}^\infty \int_0^t dt_1 \int_{t_1}^1 dt_2 \,\ldots\, \int_{t_{n-1}}^1 A(t_1)..A(t_n)\,, \tag{39}$$

where $\mathcal{P}$ denotes the path ordering symbol which orders the curve parameters from left to right according to their value beginning with the smallest one and $A(t) := A_a^j(c(t))\dot{c}^a(t)\tau_j/2$.

With this definition it is not difficult to verify the following transformation behaviour of $A(c)$ under gauge transformations and spatial diffeomorphisms respectively (recall (25)):

$$A^g(c) = g(b(c))A(c)g(f(c))^{-1} \text{ and } A^\varphi(c) = A(\varphi^{-1}(c)) \tag{40}$$

where $b(c), f(c)$ denote the beginning and final point of a curve respectively. Thus, the behaviour under gauge transformations is extremely simple which makes it easy to construct gauge invariant functions, for instance the Wilson loop functions $\mathrm{Tr}(A(c))$ where c is a closed curve, that is, a *loop*. This is the reason why QGR is also denoted as *Loop Quantum Gravity*. That holonomies also transform very naturally under spatial diffeomorphisms as depicted in the second equation of (40) has the following mathematical origin: A connection is in particular a one-form, therefore it is *naturally integrated (smeared) over one-dimensional submanifolds of* σ. Here natural means *without using a background metric*. Now the holonomy is not really the exponential of $\int_c A$ but almost as shown in (39). Thus, holonomies are precisely in accordance with our wish to construct a *background independent* quantum field theory. Moreover, the simple transformation behaviour under diffeomorphisms again makes it simple to construct spatially diffeomorphism invariant functions of holonomies: These will be functions only labelled by diffeomorphism invariance classes of loops, but these are nothing else than *knot classes*. Thus QGR has an obvious link with *topological quantum field theory (TQFT)* [32] which makes it especially attractive and was one of the major motivations for Jacobson, Rovelli and Smolin to consider Wilson loop functions for canonical quantum gravity [33]. Finally one can show [34] that the holonomies separate the points of $\mathcal{A}$, i.e. they encode all the information that is contained in a connection.

The fact that the holonomy smears A only one-dimensionally is nice due to the above reasons but it is also alarming because its functional derivative is certainly distributional and thus does not exist in an a priori mathematical sense. However, in order to obtain a well-defined Poisson algebra it is not necessary to have smooth functions of A, it is only sufficient. The key idea idea is that if we smear also the electric fields E then we might get a non-distributional Poisson algebra. By inspection from (22) it is clear that E has to be smeared in at least two dimensions in order to achieve this. Now again background independence comes to our help: Let ϵ_{abc} be the totally skew, background independent tensor density of weight -1, that is, $\epsilon_{abc} = \delta^1_{[a}\delta^2_b\delta^3_{c]}$ where [..] denotes total antisymmetrization. Then $(*E)^j_{ab} := E^j_{ab} := E^c_j\epsilon_{abc}$ is a 2-form of density weight 0. Therefore E *is naturally smeared in two dimensions*. Notice that the smearing dimensions of momenta and configuration variables add up to the dimension of σ, they are dual to each other which is a generic phenomenon for any canonical theory in any dimension. We are therefore led to consider the *electric fluxes*

$$E_j(S) = \int_S *E_j\,, \tag{41}$$

where S is a two-dimensional, open surface. It is easy to check that $E(S) := E_j(S)\tau_j$ has the following transformation behaviour

$$E^g(S) = \int_S \mathrm{Ad}_g(*E) \text{ and } E^\varphi(S) = E(\varphi^{-1}(S)). \tag{42}$$

Thus, while the transformation under spatial diffeomorphisms is again simple, the one under gauge transformations is not. However, the idea is that the $E_j(S)$ are the basic building blocks for more complicated functions of E which are already gauge invariant. The prototype of such a function is the area functional for a parameterized surface $X_S:\ D \to \sigma,\ D \subset \mathbb{R}^2$

$$\mathrm{Ar}(S) := \int_D d^2u \sqrt{\det(X_S^* q)}. \tag{43}$$

Exercise 6.
Define $n_a^S := \epsilon_{abc} X^b_{S,u^1} X^c_{S,u^2}$ and verify that (43) coincides with

$$\mathrm{Ar}(S) := \beta \int_D d^2u \sqrt{(E^a_j n^S_a)(E^b_j n^S_b)}\,, \tag{44}$$

where β is the Immirzi parameter.

It is clear that $E_j(S) = \int_D d^2u E^a_j n^S_a$ so that the area functional can be written as the limit of a Riemann sum, over small surfaces that partition S, of functions of the electric fluxes for those small surfaces.

Let us see whether the Poisson bracket between an electric flux and a holonomy is well-defined. Actually, let us be slightly more general and introduce the following notion: Let us loosely think for the moment of a graph γ as a collection of a finite number of smooth, compactly supported, oriented curves, called edges e, which intersect at most in their end points, which are called vertices v. We denote by $E(\gamma),\ V(\gamma)$ the edge and vertex set of γ respectively. A precise definition will be given in Sect. 2.1.

Definition 9.
Given a graph γ we define

$$p_\gamma:\ \mathcal{A} \to SU(2)^{|E(\gamma)|};\ A \mapsto (A(e))_{e\in E(\gamma)}. \tag{45}$$

A function $f:\ \mathcal{A} \to \mathbb{C}$ is said to be cylindrical over a graph γ, if there exists a function $f_\gamma:\ SU(2)^{|E(\gamma)|} \to \mathbb{C}$ such that $f = f_\gamma \circ p_\gamma$. We denote by $\mathrm{Cyl}^n_\gamma,\ n = 0, 1, .., \infty$ the set of n-times continuously differentiable cylindrical functions over γ and by Cyl^n the set functions which are cylindrical over *some* γ with the same differentiability type. Here we say that $f = f_\gamma \circ p_\gamma \in \mathrm{Cyl}^n_\gamma$ if and only if f_γ is n-times continuously differentiable with respect to the standard differential structure on $SU(2)^{|E(\gamma)|}$.

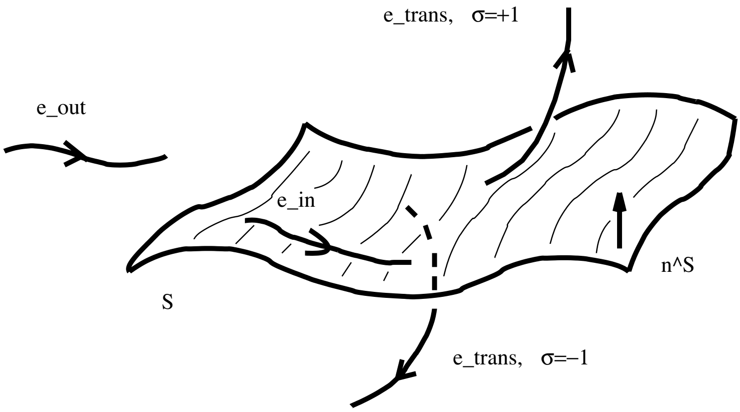


Fig. 7. Intersection structure of surfaces with edges

Our Poisson Algebra will be based on the set of functions Cyl^∞ which certainly form an Abelean Poisson subalgebra. Our next task will be to compute the Poisson bracket between a flux and an element of Cyl^∞. In order to compute this we will use the chain rule ($f \in \mathrm{Cyl}^\infty_\gamma$)

$$\{E_j(S), f\}(A) = \sum_{e\in E(\gamma)} [\frac{\partial f_\gamma(\{h_{e'}\}_{e'\in E(\gamma)})}{\partial (h_e)_{AB}} \{E_j(S), (h_e)_{AB}\}]_{|h_{e'}=A(e')} \tag{46}$$

so that the bracket will be well-defined once the bracket between a holonomy and a flux is well-defined. To compute this the intersection structure of e with S is somewhat important. In order to simplify the notation, we notice that we can always take γ to be adapted to S, that is, every edge e belongs to one of the following three types:

a) $e \in E_{\text{out}}(\gamma) \;\;\Leftrightarrow\;\; e \cap S = \emptyset$.
b) $e \in E_{\text{in}}(\gamma) \;\;\Leftrightarrow\;\; e \cap S = e$.
c) $e \in E_{\text{trans}}(\gamma) \;\;\Leftrightarrow\;\; e \cap S = b(e)$.

This can be achieved by subdividing edges into a finite number of segments and inverting their orientation if necessary as depicted in Fig. 7 (strictly speaking, this is true only if S is compactly supported, open, oriented and analytic). We also need to introduce the function $\sigma(S, e)$ which vanishes for $e \in E_{\text{in}}(\gamma) \cup E_{\text{out}}(\gamma)$ and which is ± 1 for $e \in E_{\text{trans}}(\gamma)$ if the orientations of S and e agree or disagree respectively. The easiest case is $e \in E_{\text{trans}}(\gamma),\ \sigma(S, e) = 1$. We find

$$\{E_j(S), A(e)\} = \kappa \int_D d^2u\ n_a^S(u) \int_0^1 ds\dot{e}^a(s)\delta(X_S(u), e(s))A(e_s)\frac{\tau_j}{2}A(e_s)^{-1}A(e) \tag{47}$$

where $e_s(t) := e(st)$. Noticing that the support of the δ-distribution is at $X_S(u) = e(0)$ which is an interior point of S but a boundary point of e, a careful analysis reveals that (47) reduces to

$$\{E_j(S), A(e)\} = \frac{\kappa}{4}\tau_j A(e). \tag{48}$$

With this result, (46) can be written in the compact form

$$\{E_j(S), f\}(A) = \frac{\kappa}{4} \sum_{e\in E(\gamma)} \sigma(S,e)[R_e^j f_\gamma]_{|h_{e'}=A(e')} \tag{49}$$

where we have defined the *right invariant vector fields*

$$(R_e^j f_\gamma)(\{h_{e'}\}_{e'\in E(\gamma)}\}) := (\frac{d}{dt})_{t=0} f_\gamma(\{h_{e'}\}_{e'\neq e}, e^{t\tau_j} h_e). \tag{50}$$

We can now define the vector fields v_S^j on Cyl^∞ by $v_S^j[f] := \{E_j(S), f\}$ and arrive at the Poisson *-algebra $\mathcal{B}$ generated by the v_S^j, $f \in \mathrm{Cyl}^\infty$ with involution defined by complex conjugation through the general formula $[(f,v),(f',v')] = (v[f'] - v'[f], [v,v'])$.

Exercise 7.
Fill the gaps between Eqs.(46) and (49).
Hint: Use formula 39 in order to derive (47), then expand $X_S(u) - e(t)$ around $u = u_0$ defined by $X_S(u_0) = e(0)$ and $t = 0$ to linear order in $u - u_0$ and sufficiently high order in t to arrive at (48). (Notice that e is only transversal, so $\dot{e}(0)$ may be tangential to S in $e(0)$!) Verify that the end result coincides with (49).

So we see that we arrive at a well defined algebra $\mathcal{B}$ by smearing the momenta in two dimensions. Could we also smear them in three dimensions? The answer is negative: Consider a one-parameter family of surfaces $t \mapsto S_t$ and define $E_j(\{S\}) := \int dt\, E_j(S_t)$. Then $f \mapsto \{E_j(S), f\}$ maps f out of Cyl^∞ because it involves an integral over t and thus depends on an uncountably infinite number of edges rather than a finite number. Thus this algebra would not be closed so that if we would like to stick with at least countably infinite graphs then we are *forced* to stick with two dimensional smearings of the electric fluxes!

Representation Theory of $\mathcal{B}$ and Suitable Kinematical Representations. The representation Theory of $\mathcal{B}$ has been considered only rather recently [35] and the analysis is not yet complete. However, if one sticks to irreducible representations for which 1) the flux operators are well-defined and self-adjoint (in other words, the corresponding one parameter unitary groups are weakly continuous) and 2) the representation is spatially diffeomorphism invariant, then the unique solution to the representation problem is the representation which we describe in this section.

This representation is of the form $\mathcal{H}_0 = L_2(\overline{\mathcal{A}}, d\mu_0)$ where $\overline{\mathcal{A}}$ is a certain distributional extension of $\mathcal{A}$ and μ_0 is a probability measure thereon. The most elegant description of this Hilbert space uses the theory of C^*-algebras [36] but fortunately there is a purely geometric description available [37] which is easier to access for the beginner. In what follows we assume for simplicity that σ is an oriented, connected, simply connected smooth manifold. One can show that each smooth manifold admits at least one analytic structure (i.e. the atlas of charts consists of real analytic maps) and we assume to have picked one once and for all.

Curves, Paths, Graphs, and Groupoids.

DEFINITION 10.

i) By a curve c we mean a map $c: \; [0,1] \to \sigma$ which is piecewise analytic, continuous, oriented and an embedding (does not come arbitrarily close to itself). It is automatically compactly supported. The set of curves is denoted $\mathcal{C}$ in what follows.
ii) On $\mathcal{C}$ we define maps $\circ$, $(.)^{-1}$ called composition and inversion respectively by

$$[c_1 \circ c_2](t) = \begin{cases} c_1(2t) & \text{for } \; t \in [0, \frac{1}{2}] \\ c_2(2t-1) & \text{for } \; t \in [\frac{1}{2}, 1] \end{cases} \tag{51}$$

if $f(c_1) = b(c_2)$ and

$$c^{-1}(t) = c(1-2t). \tag{52}$$

iii) By a path p we mean an equivalence class of curves c which differ from each other by a finite number of reparameterizations and retracings, that is, $c \sim c'$ if there either exists a map $t \mapsto f(t), \; \dot{f}(t) > 0$ with $c = c' \circ f$ or we may write c, c' as compositions of segments in the form $c = s_1 \circ s_2, \; c' = s_1 \circ s_3 \circ s_3^{-1} \circ s_2$ (and finite combinations of such moves). Notice that a curve induces its orientation and its end points on its corresponding path. The set of paths is denoted by $\mathcal{P}$.
iv) By a graph γ we mean a finite collection of elements of $\mathcal{P}$. We may break paths into pieces such that γ can be thought of as a collection of edges $e \in E(\gamma)$, that is, paths which have an entire analytic representative and which intersect at most in their end points $v \in V(\gamma)$ called vertices. The set of graphs is denoted by Γ.

These objects are depicted in Fig. 8.

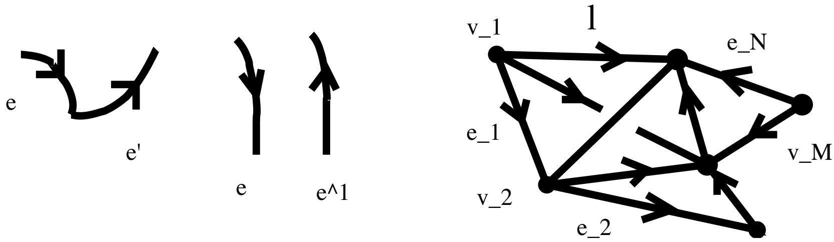


Fig. 8. Paths and graphs

Exercise 8.

a) Despite the name, composition and inversion does not equip $\mathcal{C}$ with a group structure for many reasons. Verify that composition is not associative and that the curve $c \circ c^{-1}$ is not simply $b(c)$ but rather a retracing. Moreover, contemplate that $\mathcal{C}$ does not have a unit and that not every two elements can be composed.
b) Define composition and inversion of paths by taking the equivalence class of the compositions and inversions of any of their representatives and check that this definition is well defined. Check that then composition of paths is associative and that $p \circ p^{-1} = b(p)$. However, $\mathcal{P}$ still does not have a unit and still not every two elements can be composed.
c) Let Obj:= σ and for each $x, y \in \sigma$ let $\mathrm{Mor}(x, y) := \{p \in \mathcal{P} :\ b(p) = x,\ f(p) = y\}$. Recall the mathematical definition of a category and conclude that $\mathcal{P}$ is a category in which every morphism is invertible, that is, a *groupoid*.
d) Define the relation $\prec$ on Γ by saying that $\gamma \prec \gamma'$ if and only if every $e \in E(\gamma)$ is a finite composition of the $e' \in E(\gamma')$ and their inverses. Verify that $\prec$ equips Γ with the structure of a directed set, that is, for each $\gamma, \gamma' \in \Gamma$ we find $\gamma'' \in \Gamma$ such that $\gamma, \gamma' \prec \gamma''$.
Hint: For this to work, analyticity of the curve representatives is crucial. Smooth curves can intersect in Cantor sets and thus define graphs which are no longer finitely generated. Show first that this is not possible for analytic curves.
e) Given a curve c with path equivalence class p; notice that for the holonomy with respect to $A \in \mathcal{A}$ holds $A(c) = A(p)$. Contemplate that, in particular, every group is a groupoid and that every connection $A \in \mathcal{A}$ qualifies as a groupoid homomorphism, that is, $A :\ \mathcal{P} \to SU(2);\ p \mapsto A(p)$ with

$$A(p \circ p') = A(p)A(p') \text{ and } A(p^{-1}) = (A(p))^{-1} . \tag{53}$$

The fact that holonomies are really defined on paths rather than curves and that holonomies are characterized algebraically by 53 makes the following definition rather natural.

Definition 11.
The quantum configuration space is defined as the set $\overline{\mathcal{A}} := \mathrm{Hom}(\mathcal{P}, SU(2))$ of all algebraic, arbitrarily non-continuous groupoid morphisms.

Here non-continuous means that in contrast to $A \in \mathcal{A}$ for an element $A \in \overline{\mathcal{A}}$ it is possible that $A(p) = 1$ varies discontinuously as we vary p continuously. Thus, $\overline{\mathcal{A}}$ can be thought of as a distributional extension of $\mathcal{A}$.

Topology on $\overline{\mathcal{A}}$. So far $\overline{\mathcal{A}}$ is just a set. We now equip it with a topology. The idea is actually quite simple. Recall the maps (45) which easily extend from $\mathcal{A}$ to $\overline{\mathcal{A}}$ and maps $\overline{\mathcal{A}}$ into $SU(2)^{|E(\gamma)|}$. Now $SU(2)^{|E(\gamma)|}$ is a compact Hausdorff topological group[5] in its natural manifold topology and we would like to exploit that. Thus we are motivated to consider the spaces $X_\gamma := \mathrm{Hom}(\gamma, SU(2)^{|E(\gamma)|})$ where γ is considered as a subgroupoid of Γ with objects $V(\gamma)$ and morphisms generated by the $e \in E(\gamma)$. The map

$$X_\gamma \to SU(2)^{|E(\gamma)|};\ x_\gamma \mapsto \{x_\gamma(e)\}_{e\in E(\gamma)} \tag{54}$$

identifies X_γ with $G^{|E(\gamma)|}$ since $x_\gamma \in X_\gamma$ is already defined by which values it takes on the $e \in E(\gamma)$ and we may thus use this identification in order to equip X_γ with a compact Hausdorff topology. Now consider the uncountably infinite product

$$X_\infty := \prod_{\gamma\in\Gamma} X_\gamma \tag{55}$$

A standard result from general topology, Tychonov's theorem, tells us that the smallest topology on X_∞ such that all the maps $p_\gamma :\ X_\infty \to X_\gamma;\ (x_\gamma)_{\gamma\in\Gamma} \mapsto x_\gamma$ are continuous is a compact Hausdorff topology[6]. Now we would like to identify $\overline{\mathcal{A}}$ with X_∞ through the restriction map

$$\Phi' :\ \overline{\mathcal{A}} \to X_\infty;\ A \mapsto (x_\gamma := A_{|\gamma} = p_\gamma(A))_{\gamma\in\Gamma} \tag{56}$$

However, that map cannot be surjective because the points of $\overline{\mathcal{A}}$ satisfy the following constraint which encodes the algebraic properties of a generalized connection: Let $\gamma \prec \gamma'$ and define the graph restriction maps

$$p_{\gamma'\gamma} :\ X_{\gamma'} \to X_\gamma;\ x_{\gamma'} \mapsto (x_{\gamma'})_{|\gamma} \tag{57}$$

which satisfy the compatibility condition

$$p_{\gamma''\gamma} = p_{\gamma'\gamma} \circ p_{\gamma''\gamma'} \text{ for } \gamma \prec \gamma' \prec \gamma'' \tag{58}$$

Then automatically

$$p_{\gamma'\gamma}(A_{|\gamma'}) = A_{|\gamma} \tag{59}$$

We are therefore forced to consider the subset of X_∞ defined by

$$\overline{X} := \{(x_\gamma)_{\gamma\in\Gamma} \in X_\infty;\ p_{\gamma'\gamma}(x'_\gamma) = x_\gamma\ \forall\, \gamma \prec \gamma'\} \tag{60}$$

---

[5] Here it is crucial that $G = SU(2)$ is compact and thus for non-real Immirzi parameter all of what follows would be false.

[6] Recall that we know the topology on a space when we know a base of open sets from which we obtain all open sets by arbitrary unions and finite intersections. Since the preimages of open sets under continuous functions are open by definition, we obtain a topology once we know which functions are continuous.

Exercise 9.

i) Show that the maps (57) are continuous surjections.
Hint: Exploit the identification of the X_γ with powers of $SU(2)$ and the continuity of multiplication and inversion in groups to establish continuity. To establish surjectivity use the fact that each edge e of γ contains an edge e'_e of γ' as a segment such that the e'_e do not overlap each other. Now given x_γ set $x_{\gamma'}(e'_e) = x_\gamma(e)$ and extend trivially away from the e'_e. Check that this defines an element of $X_{\gamma'}$.

ii) Show that $\overline{X}$ is a closed subset of X_∞.
Hint: Since $\overline{X}$ is not a metric space we must work with nets and show that every net of points $x^\alpha \in \overline{X}$ which converges in X_∞ actually converges in $\overline{X}$. Using the definition of the topology on X_∞, show that this is equivalent to showing that the $p_\gamma(x^\alpha) = x^\alpha_\gamma$ converge to points x_γ which satisfy (59) and verify this using continuity of the $p_{\gamma'\gamma}$ just established.

The surjectivity of the $p_{\gamma'\gamma}$ qualifies $\overline{X}$ as the so-called projective limit of the X_γ, a mathematical structure which is independent of our concrete context once we have a directed index set Γ at our disposal and surjective projections which satisfy the compatibility condition (58).

Now another standard result from topology now tells us that $\overline{X}$, being the closed subset of a compact Hausdorff space, is a compact Hausdorff space in the subspace topology and the question arises whether

$$\Phi : \overline{\mathcal{A}} \to \overline{X};\ A \mapsto (x_\gamma := A_{|\gamma} = p_\gamma(A))_{\gamma\in\Gamma} \tag{61}$$

is a bijection. Injectivity is fairly easy to see while surjectivity is a little bit tricky.

Exercise 10.
Show that (61) is a bijection.
Hint: Given $x \in \overline{X}$ and $p \in \mathcal{P}$ choose any $\gamma_p \in \Gamma$ such that $p \in \gamma_p$ and define A_x by $A_x(p) := x_{\gamma_p}(p)$. Show that this definition is well defined using the directedness of Γ and that A_x is a groupoid homomorphism.

Let us collect these results in the following theorem [38].

Theorem 3.
The space $\overline{\mathcal{A}}$ equipped with the weakest topology such that the maps p_γ of (45) are continuous, is a compact Hausdorff space.

The value of this result is that it gives us a powerful tool for constructing measures on $\overline{\mathcal{A}}$.

Measures on $\overline{\mathcal{A}}$. A powerful theorem due to Riesz and Markov, sometimes called the Riesz representation theorem, tells us that there is a one-to-one correspondence between the positive linear functionals Λ on the algebra $C(\overline{\mathcal{A}})$ of continuous functions on a compact Hausdorff space $\overline{\mathcal{A}}$ and (regular, Borel) probability measures μ thereon through the simple formula

$$\Lambda(f) := \int_{\overline{\mathcal{A}}} d\mu(A)\ f(A) \tag{62}$$

One says Λ is represented by f. Here a linear functional is called positive if $\Lambda(|f|^2) \geq 0$ for any $f \in C(\overline{\mathcal{A}})$. A function algebra on a compact space can be equipped with the sup-norm $||f|| := \sup_{A \in \overline{\mathcal{A}}} |f(A)|$ which evidently has the so-called C^*-property $||f\overline{f}|| = ||f||^2$ so that (w.l.g. we may take $C(\overline{\mathcal{A}})$ to be complete w.r.t. the norm) $C(\overline{\mathcal{A}})$ is a C^*-algebra. A standard result in functional analysis reveals that positive linear functionals on C^*-algebras are automatically continuous, $|\Lambda(f)| \leq \Lambda(1)\ ||f||$ and if we choose the normalization of Λ to be $\Lambda(1) = 1$ then μ is a probability measure.

In order to specify the measure μ_0 that we are interested in, it is therefore enough to specify a positive linear functional Λ_0. The most elegant way of defining Λ_0 is through the following definition.

Definition 12.

i) Given a graph γ, label each edge $e \in E(\gamma)$ with a triple of numbers (j_e, m_e, n_e) where $j_e \in \{\frac{1}{2}, 1, \frac{3}{2}, 2, ..\}$ is a half-integral spin quantum number and $m_e, n_e \in \{-j_e, -j_e+1, .., j_e\}$ are magnetic quantum numbers. A quadruple

$$s := (\gamma, \boldsymbol{j} := \{j_e\}_{e \in E(\gamma)}, \boldsymbol{m} := \{m_e\}_{e \in E(\gamma)}, \boldsymbol{n} := \{n_e\}_{e \in E(\gamma)}) \tag{63}$$

is called a spin network (SNW). We also write $\gamma(s)$ etc. for the entries of a SNW.

ii) Choose once and for all one representative ρ_j, $j > 0$ half integral, from each equivalence class of irreducible representations of $SU(2)$. Then

$$T_s : \overline{\mathcal{A}} \to \mathbb{C};\ A \mapsto \prod_{e \in E(\gamma)} [\sqrt{2j_e+1}[\rho_{j_e}(A(e))]_{m_e n_e}] \tag{64}$$

is called the spin-network function (SNWF) of s. Here $[\rho_j(.)]_{mn}$ denotes the matrix elements of the matrix valued function $\rho_j(.)$.

An example of a SNW, which can be arbitrarily large and with vertices of arbitrarily high valence, is given in Fig. 9. The original motivation for the definition of spin network functions [40] in loop quantum gravity was the fact that they are linearly independent in contrast to the Wilson loop functions which suffer from the so-called *Mandelstam identities*. For $SU(2)$ matrices h, h' they are $\mathrm{Tr}(h)\ \mathrm{Tr}(h') = \mathrm{Tr}(hh') + \mathrm{Tr}(h(h')^{-1})$ and $\mathrm{Tr}(h) = \mathrm{Tr}(h^{-1})$ which leads to an infinite tower of identities of the form

$$[\mathrm{Tr}(A(\alpha_1))\ \mathrm{Tr}(A(\alpha_2))]\mathrm{Tr}(A(\alpha_1)) = \mathrm{Tr}(A(\alpha_1))[\mathrm{Tr}(A(\alpha_2))\ \mathrm{Tr}(A(\alpha_1))] \tag{65}$$

depending on how we bracket the product of traces involving the three loops $\alpha_1, \alpha_2, \alpha_3$ with a common base point. The SNWF's remove these cumbersome identities first by labelling functions by edges rather than loops and secondly by the simple observation that a tensor product of (fundamental) representations can be uniquely decomposed into irreducibles (Clebsh-Gordon decomposition).

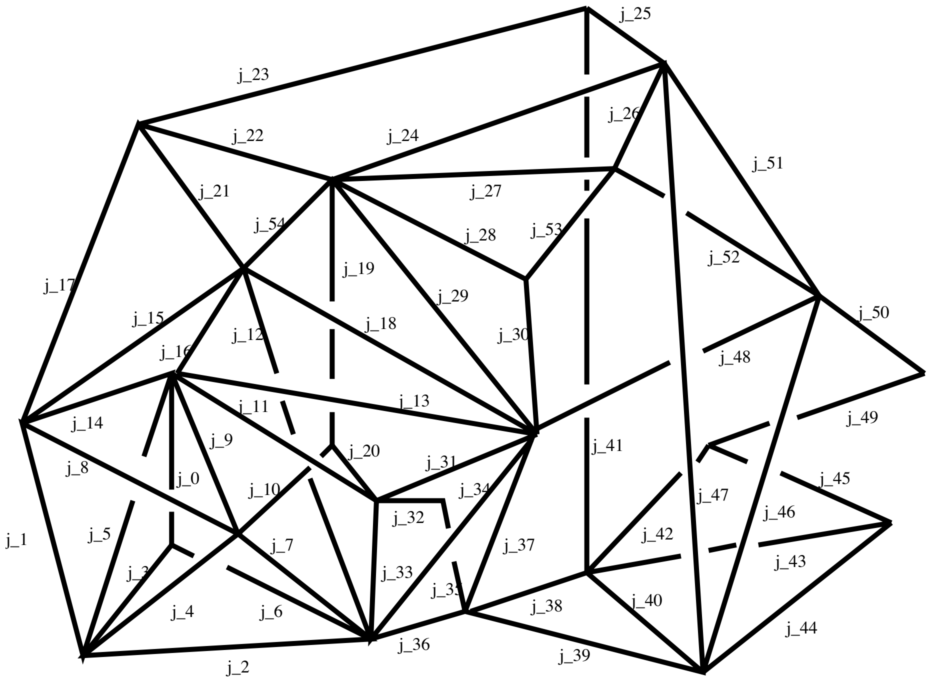


Fig. 9. A SNW. Orientations and magnetic quantum numbers are suppressed

Theorem 4.
The uniform (Ashtekar–Lewandowski) measure μ_0 is uniquely defined by the positive linear functional [39]

$$\Lambda_0(T_s) := \begin{cases} 1 & \text{for } s = (\emptyset, \mathbf{0}, \mathbf{0}, \mathbf{0}) \\ 0 & \text{otherwise} \end{cases} \tag{66}$$

Exercise 11.

i) Recall the representation theory of $SU(2)$ from the quantum mechanics of angular momentum and verify that the SNWF are indeed linearly independent.

ii) Verify that Λ_0 is a positive linear functional.
Hint: Using the Stone–Weierstrass theorem, show first that the finite linear combinations of SNWF are dense in $C(\overline{\mathcal{A}})$. By continuity of Λ_0 it is therefore sufficient to check positivity on finite linear combinations

$$f = \sum_{n=1}^{N} z_n T_{s_n}, \quad N < \infty, \ z_n \in \mathbb{C} \tag{67}$$

with s_n mutually different SNW's. To see this, verify that $\Lambda_0(\overline{T_s} T_{s'}) = 0$ for $s \neq s'$ by using the Clebsh–Gordon formula $j \otimes j' \equiv (j+j') \oplus (j+j'-1) \oplus .. \oplus (|j-j'|)$.

iii) A fundamental theorem for the representation theory of compact groups is due to Peter and Weyl [41]. For $SU(2)$ it amounts to saying that the functions

$$T_{jmn} : \ SU(2) \to \mathbb{C}; \ h \mapsto \sqrt{2j+1}[\rho_j(h)]_{mn} \tag{68}$$

form an orthonormal basis for the Hilbert space $L_2(SU(2), d\mu_H)$ where μ_H is the normalized Haar measure on $SU(2)$ (the unique normalized measure which invariant under inversion as well as left and right translation in $SU(2)$). Based on this result, show that the SNWF form an orthonormal basis for the Hilbert space $L_2(\overline{\mathcal{A}}, d\mu_0)$.

Let us summarize the results of the exercise in the following theorem [40].

THEOREM 5.
The kinematical Hilbert space $\mathcal{H}_{\text{kin}} := L_2(\overline{\mathcal{A}}, d\mu_0)$ *defined by (66) is non-separable and has the SNWF's* T_s *as orthonormal basis.*

Representation Property. So far we did not verify that $\mathcal{H}_{\text{kin}}$ is a representation space for our *-algebra $\mathcal{B}$ of basic operators. This will be done in the present section. Indeed, until today no other irreducible representation of the holonomy-flux algebra has been found (except if one allows also infinite graphs [42]).

By theorem (6) the subspace of finite linear combinations of SNWF's is dense in $\mathcal{H}_{\text{kin}}$ with respect to the L_2 norm. On the other hand, we notice that the definition of $\text{Cyl}^\infty(\mathcal{A})$ simply extends to $\text{Cyl}^\infty(\overline{\mathcal{A}})$ and that finite linear combinations of SNWF's form a subspace of $\text{Cyl}^\infty(\overline{\mathcal{A}})$. Thus, we may choose $\mathcal{D}_{\text{kin}} := \text{Cyl}^\infty(\overline{\mathcal{A}})$ and obtain a dense, invariant domain of $\mathcal{B}$ as we will see shortly. We define a representation of the holonomy-flux algebra by $(f' \in \text{Cyl}^\infty(\mathcal{A}),\ f \in \text{Cyl}^\infty(\overline{\mathcal{A}}),\ A \in \overline{\mathcal{A}})$

$$\begin{aligned}
[\pi(f) \cdot f'](A) &:= (f'f)(A) \\
[\pi(v_S^j) \cdot f](A) &= [\pi(v_S^j)\pi(f) \cdot 1](A) = [([\pi(v_S^j), \pi(f)] + \pi(f)\pi(E_j(S))) \cdot 1](A) \\
&:= i\hbar[\pi(v_S^j[f]) \cdot 1](A) = i\hbar(v_S^j[f])(A).
\end{aligned} \tag{69}$$

Thus $\pi(f)$ is a multiplication operator while $\pi(v_S^j)$ is a true derivative operator, i.e. it annihilates constants. Notice that the canonical commutation relations are already obeyed by construction, thus we only need to verify the *-relations and the fact that $\pi(v_S^j)$ annihilates constants will be crucial for that.

The $\pi(f)$ are bounded multiplication operators (recall that smooth, i.e. in particular continuous, functions on compact spaces are uniformly bounded, that is, have a sup-norm) so that the adjoint is complex conjugation, therefore there is nothing to check. As for $\pi(v_S^j)$ we notice that given two smooth cylindrical functions on $\overline{\mathcal{A}}$ we always find a graph γ over which both of them are cylindrical and which is already adapted to S.

EXERCISE 12.
Let f be cylindrical over γ. Verify that

$$\Lambda_0(f) = \int_{SU(2)^{|E(\gamma)|}} \prod_{e \in E(\gamma)} d\mu_H(h_e) f_\gamma(\{h_e\}_{e \in E(\gamma)}). \tag{70}$$

Hint: Write f as a (Cauchy limit of) finite linear combinations of SNWF's and verify that (70) coincides with (66).

Using the explicit expression (49) and the result of exercise 12 it is easy to see that the symmetry condition $< f, \pi(v^j_S) f' >_{\rm kin} = < \pi(v^j_S) f, f' >_{\rm kin}$ is equivalent with the condition

$$< F, R^j F' >_{L_2(SU(2), d\mu_H)} = - < R^j F, F' >_{L_2(SU(2), d\mu_H)} \tag{71}$$

for any $F, F' \in C^\infty(SU(2))$ and R^j is the right invariant vector field on $SU(2)$. However, μ_H is by definition invariant under left translations and R^j is a generator of left translations in $SU(2)$ so the result follows. This shows that $\mathcal{D}_{\rm kin}$ is contained in the domain of $\pi(v^j_S)^\dagger$ and that the restriction of the adjoint to $\mathcal{D}_{\rm kin}$ coincides with $\pi(v^j_S)$. That $\mathcal{D}_{\rm kin}$ is actually a domain of (essential) self-adjointness requires a little bit more work but is not difficult to see, e.g. [1].

Finally, let us verify that the representation is irreducible. By definition, a representation is irreducible if every non-zero vector is cyclic and a vector Ω is cyclic if the set of vectors $\pi(a)\Omega,\ a \in \mathcal{B}$ is dense. Now the vector $\Omega = 1$ is cyclic because the vectors $\pi(f)\Omega = f,\ f \in {\rm Cyl}^\infty$ are already dense. Given an arbitrary element $\psi \in \mathcal{H}_{\rm kin}$ we know that it is a Cauchy limit of finite linear combinations of spin network functions. Thus, if we can show that we find a sequence $a_n \in \mathcal{B}$ such that $\pi(a_n)\psi$ converges to Ω, then we are done. It is easy to see (exercise) that this problem is equivalent to showing that any $F \in L_2(F, d\mu_H)$ can be mapped by the algebra formed out of right invariant vector fields and smooth functions on $SU(2)$ to the constant function.

Exercise 13.
Check that this is indeed the case.
Hint: Show first that it is sufficient to establish that any polynomial p of the $a, b, c, d,\ ad - bc = 1$ for $h = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(2)$ can be mapped to the constant function. Show then that suitable linear combinations of the $R^j,\ j = 1, 2, 3$ with coefficients in $C^\infty(SU(2))$ produce the derivatives $\partial_a, \partial_b, \partial_c$ and convince yourself that $a^N p$ is a polynomial in a, b, c for sufficiently large N.

Let us collect these results in the following theorem [43].

Theorem 6.
The relations (69) define an irreducible representation of $\mathcal{B}$ on $\mathcal{H}_{\rm kin}$.

Thus, the representation space $\mathcal{H}_{\rm kin}$ constructed satisfies all the requirements that qualify it as a good kinematical starting point for solving the quantum constraints. Moreover, the measure μ_0 is spatially diffeomorphism invariant as we will see shortly and together with the uniqueness result quoted at the beginning of this section, this is the only representation with that property. There are, however, doubts on physical grounds whether one should insist on a spatially diffeomorphism invariant representation because the smooth and even analytic structure of σ which is encoded in the spatial diffeomorphism group should not play a fundamental role at short scales if Planck scale physics is fundamentally discrete. In fact, as we will see later, QGR predicts a discrete Planck scale structure and therefore the fact that we started with analytic data and ended up with discrete (discontinuous) spectra of operators looks awkward. Therefore, on the one hand, we should keep in mind that other representations are possibly better suited in the final picture; on the other hand, there is no logical contradiction within the present formulation and in fact in 2+1 gravity one has a final combinatorial description while one started with analytical structures as well.

2.2 Quantum Kinematics

In this section we discuss the complete solution of the Gauss and Vector constraint as well as the quantization of kinematical, geometrical operators that measure the length, area and volume of *coordinate* curves, surfaces and regions respectively. We call these results *kinematical* because the Gauss and Vector constraint do not generate dynamics, this is the role of the Hamiltonian constraint which we will discuss in the third part. Moreover, the kinematical geometrical operators do not commute with the Vector constraint or the Hamiltonian constraint and are therefore not Dirac observables. However, as we will show, one can turn these operators easily into Dirac observables, at least with respect to the Vector constraint, and the fact that the spectrum is discrete is robust under those changes.

The Space of Solutions to the Gauss and Spatial Diffeomorphism Constraint. Recall the transformation behaviour of classical connections $A \in \mathcal{A}$ under $SU(2)$ gauge transformations and spatial diffeomorphisms (40). These equations trivially lift from $\mathcal{A}$ to $\overline{\mathcal{A}}$ and we may construct corresponding operators as follows: Let $\overline{\mathcal{G}} := \mathrm{Fun}(\Sigma, SU(2))$ be the set of local gauge transformations without continuity requirement and consider the set $\mathrm{Diff}^\omega(\Sigma)$ of *analytic* diffeomorphisms. We are forced to consider analytic diffeomorphisms as otherwise we would destroy the analyticity of the elements of Γ. These two groups have a natural semi-direct product structure that has its origin in the algebra (23) and is given by

$$[\overline{\mathcal{G}} \rtimes \mathrm{Diff}^\omega(\Sigma)] \times [\overline{\mathcal{G}} \rtimes \mathrm{Diff}^\omega(\Sigma)] \to [\overline{\mathcal{G}} \rtimes \mathrm{Diff}^\omega(\Sigma)] \tag{72}$$
$$[g, \varphi] \cdot [g', \varphi'] = [g(g' \circ \varphi^{-1}), \varphi \circ \varphi'] \,. \tag{73}$$

EXERCISE 14.
Verify (72).
Hint: Define $[g, \mathrm{id}] \cdot A := A^g$, $[\mathrm{id}, \varphi] \cdot A := A^\varphi$ and $[g, \varphi] \cdot A := [g, \mathrm{id}] \cdot ([\mathrm{id}, \varphi] \cdot A)$.

We now define representations

$$\begin{aligned} \hat{U} :\ & \overline{\mathcal{G}} \to \mathcal{B}(\mathcal{H}_{\mathrm{kin}});\ g \mapsto \hat{U}(g) \\ \hat{V} :\ & \mathrm{Diff}^\omega(\Sigma) \to \mathcal{B}(\mathcal{H}_{\mathrm{kin}});\ \varphi \mapsto \hat{V}(\varphi) \end{aligned} \tag{74}$$

densely on $f = p_\gamma^* f_\gamma \in \mathcal{D}_{\mathrm{kin}}$ by

$$\begin{aligned} [\hat{U}(g) f](A) &:= f_\gamma(\{g(b(e))\ A(e)\ g(f(e))^{-1}\}_{e \in E(\gamma)}) \\ [\hat{V}(\varphi) f](A) &:= f_\gamma(\{A(\varphi^{-1}(e))\}_{e \in E(\gamma)}) \,. \end{aligned} \tag{75}$$

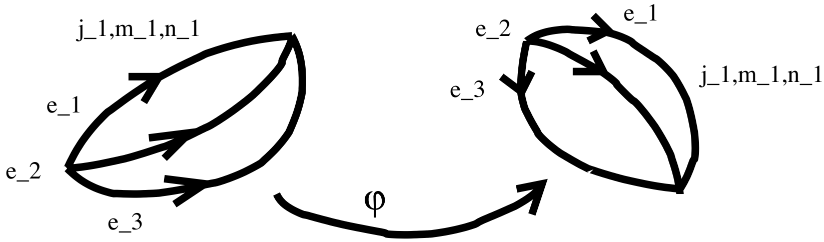


Fig. 10. Action of Spatial Diffeomorphisms on SNW's

Here $\mathcal{B}(.)$ denotes the bounded operators on a Hilbert space. This definition of course comes precisely from the classical formula (40). The action of a diffeomorphism on a SNWF T_s is therefore simply by mapping the graph $\gamma(s)$ to $\varphi^{-1}(s)$ while the labels j_e, m_e, n_e are carried from e to $\varphi^{-1}(e)$ as depicted in Fig. 10.

Then the following theorem holds [43].

THEOREM 7.
The relations (75) define a unitary representation of the semi-direct product kinematical group $\overline{\mathcal{G}} \rtimes \mathrm{Diff}^{\omega}(\Sigma)$.

EXERCISE 15.
Prove theorem (7).
Hint: Check unitarity on the SNWF basis using the bi-invariance of the Haar measure. That (75) holds can be traced back to exercise 14.

The unitarity of the kinematical gauge group implies invariance of the measure μ_0 and thus supplies additional motivation for the representation space $\mathcal{H}_{\mathrm{kin}}$. Notice that the statement that (75) defines a representation in particular means that the kinematical constraint algebra is *free of anomalies*. This should be contrasted with string theory where the anomaly sits also in the spatial diffeomorphism group (e.g. $\mathrm{Diff}(S^1)$ for the closed string) unless one chooses the critical dimension $D = 25(9)$ for the bosonic (supersymmetric) string.

Let us now solve the kinematical constraints. By definition, we are supposed to find algebraic distributions $l \in \mathcal{D}^*_{\mathrm{kin}}$ which satisfy

$$l(\hat{U}(g)f) = l(\hat{V}(\varphi)f) = l(f) \;\; \forall\, g \in \overline{\mathcal{G}},\; \varphi \in \mathrm{Diff}^{\omega}(\Sigma),\; f \in \mathcal{D}_{\mathrm{kin}}\,. \tag{76}$$

Now it is not difficult to see that any element of $\mathcal{D}^*_{\mathrm{kin}}$ can be conveniently written in the form

$$l(.) = \sum_s \; c_s \;\; < T_s, . >_{\mathrm{kin}}\,, \tag{77}$$

where c_s are complex valued coefficients and the uncountably infinite sum extends over all possible SNW's. The general solution to (76) is then easy to describe: Invariance under $\overline{\mathcal{G}}$ means that for fixed γ the coefficients $c_{\gamma,\boldsymbol{j},\boldsymbol{m},\boldsymbol{n}}$ have to be chosen, as $\boldsymbol{j}, \boldsymbol{m}, \boldsymbol{n}$ vary, in such a way that at each vertex of γ the resulting function is gauge invariant. That is, if $j_1, .., j_n$ are the labels of edges incident at v, then the c_s have to arrange themselves to a projector on the trivial representations contained in the tensor product $j_1 \otimes .. \otimes j_n$. Such a projector is also called intertwiner in the mathematical literature. For $SU(2)$ this leads to the theory of Clebsh-Gordon coefficients, $6j$-symbols etc. As for $\mathrm{Diff}^\omega(\Sigma)$ we see that $c_{\varphi(\gamma),\boldsymbol{j},\boldsymbol{m},\boldsymbol{n}}$ must be independent of φ, therefore $c_{\varphi,\boldsymbol{j},\boldsymbol{m},\boldsymbol{n}}$ depends only on the *generalized knot class of* γ! We say generalized because, as we will see later on, the physically relevant graphs are those with self-intersections while classical knot theory deals only with smooth curves.

One may ask whether one should already define a physical inner product with respect to the Gauss and spatial Diffeomorphism constraint and then solve the Hamiltonian constraint in a second, separate step on that already partly physical Hilbert space . While such a Hilbert space can indeed be constructed [43] it is of no use for QGR because the Hamiltonian constraint cannot leave that Hilbert space invariant as we see from the second equation in (14) and we must construct the physical inner product from the full solution space to all constraints. However, at least with respect to the kinematical constraints the full quantization programme including the question of observables has already been completed except for the analysis of the classical limit.

Kinematical Geometrical Operators. We will restrict ourselves to the description of the area operator the classical expression of which we already wrote in (43) and (44).

In order to quantize $\mathrm{Ar}(S)$ one starts from (44) and decomposes the analytical, compactly supported and oriented surface S or, equivalently, its preimage D under X_S into small pieces S_I. Then the exact area functional is approximated by the Riemann sum

$$\mathrm{Ar}(\{S\}) = \beta \sum_I \sqrt{E_j(S_I)^2}\,. \tag{78}$$

This function is easily quantized because $\hat{E}_j(S_I) = i\hbar v_s^j$ is a self-adjoint operator so that the sum over j of its squares is positive semi-definite, hence its square root is well-defined. Let us denote the resulting, partition dependent operator by $\widehat{\mathrm{Ar}}(\{S\})$. Now one can show that the (strong) limit as the partition is sent to the continuum exists [44] and a partition independent operator $\widehat{\mathrm{Ar}}(S)$ results [44].

THEOREM 8.
The area functional admits a well-defined quantization $\widehat{\mathrm{Ar}}(S)$ on $\mathcal{H}_{\mathrm{kin}}$ with the following properties:

i) $\widehat{\mathrm{Ar}}(S)$ is positive semidefinite, (essentially) self-adjoint with $Cyl^2(\overline{\mathcal{A}})$ as domain of (essential) self-adjointness.
ii) The spectrum $Spec(\widehat{\mathrm{Ar}}(S))$ is pure point (discrete) with eigenvectors being given by finite linear combinations of spin network functions.

iii) *The eigenvalues are given explicitly by*

$$\begin{aligned}\lambda_{j_1,j_2,j_{12}} &= \frac{\beta \ell_{\mathrm{P}}^2}{4}\sqrt{2j_1(j_1+1)+2j_2(j_2+1)-j_{12}(j_{12}+1)} \\ j_{12} &\in \{j_1+j_2, j_1+j_2-1, .., |j_1-j_2|\}\,,\end{aligned} \tag{79}$$

where j_1, j_2 *are spin quantum numbers and* $\ell_{\mathrm{P}}^2 = \hbar\kappa$ *is the Planck area. The spectrum has an area gap (smallest non-vanishing eigenvalue) given by*

$$\lambda_0 = \beta \ell_{\mathrm{P}}^2 \frac{\sqrt{3}}{4}\,. \tag{80}$$

iv) *Spec*$(\widehat{\mathrm{Ar}}(S))$ *contains information about the topology of* S*, for instance it matters whether* $\partial S = \emptyset$ *or not.*

EXERCISE 16.
Verify that the area gap is indeed given by (80) and check that the distance between subsequent eigenvalues rapidly decreases as $j_1, j_2 \to \infty$. Can one give an asymptotic formula for $N(A)$, the number of eigenvalues (discarding multiplicity) in the interval $[A - \ell_{\mathrm{P}}^2, A + \ell_{\mathrm{P}}^2]$? Thus, a correspondence principle, important for the classical limit is valid. If the spectrum would only consist of the *main series* $\lambda_j = \frac{\ell_{\mathrm{P}}^2}{2}\sqrt{j(j+1)}$ which one obtains for $j_1 = j_2 = j$, $j_{12} = 0$ then such a correspondence principle would certainly not hold which is, e.g., relevant for the black body spectrum of the Hawking radiation.

Theorem 8 is an amazing result for several reasons:

A) First of all, the expression for $\mathrm{Ar}(S)$ depends non-polynomially, not even analytically on the product $E^a_j(x)E^b_j(x)$, $x \in S$. Now $E^a_j(x)$ becomes an operator valued distribution in the quantum theory and products of distributions at the same point are usually badly divergent. However, $\widehat{\mathrm{Ar}}(S)$ is *perfectly well-defined!* This is the first pay-off for sticking to a rigorous and background independent formalism!

B) Although $S, \gamma, \Sigma, ..$ are analytical, the spectrum $\mathrm{Spec}(\widehat{\mathrm{Ar}}(S))$ is discrete. In other words, suppose we are measuring the area of a sheet of paper with a spin-network state. As long as the sheet does not cut an edge of the graph, the area eigenvalue is exactly zero *no matter how "close" the edge and the sheet are.* We have put the word "close" in inverted commas because this word has no meaning: Since there is no background metric, we do not know what close means, only diffeomorphism invariant notions have a meaning such as "the edge is cut" or "the edge is not cut". However, once the edge is cut the area eigenvalue jumps at least by the area gap. This strongly hints that the microscopical geometry is really distributional (discontinuous) and that we have a discrete Planck scale structure, the role of the *atoms of geometry* being played by the one-dimensional (polymer-like) excitations labelled by SNW's. One may speculate that this discrete structure is fundamental and that the analyticity assumptions that we began with should be unimportant, in the final picture everything should be only combinatorial. The smooth geometry that we are familiar with at macroscopic scales is merely a result of coarse graining, for instance in order that a SNWF labelled with spin $j = 1/2$ on every edge assigns to our sheet of paper its area of about 100cm^2, an order of 10^{68} edges of the SNW have to cut the sheet!

C) Qualitatively similar results apply to the volume operator $\widehat{\text{Vol}}(R)$ [44,45] and the length operator $\widehat{\text{Len}}(c)$ [46] whose classical expressions are given by

$$\text{Vol}(R) = \int_R d^3x \sqrt{\det(q)} \text{ and } \text{Len}(c) = \int_c \sqrt{q_{ab} dx^a dx^b}. \tag{81}$$

D) These kinematical operators are certainly not Dirac observables because they are not even spatially diffeomorphism invariant (but $SU(2)$ invariant) since the objects R, S, c are just coordinate submanifolds of Σ. Thus, one may wonder whether the properties of the spectrum just stated have any significance at all. The answer is believed to be affirmative as the following argument shows: For instance, instead of $\text{Vol}(R)$ consider

$$\text{Vol}_{EM} = \int_\Sigma d^3x \sqrt{\det(q)} \;\; \theta \left(\frac{q_{ab}}{\sqrt{det(q)}} [E^a E^b + B^a B^b] \right) , \tag{82}$$

where we have coupled a Maxwell field to GR with electromagnetic fields E^a, E^b and θ is the step function. The physical meaning of (82) is that it measures the volume of the region where the electromagnetic field energy density is non-vanishing and it is easy to check that (82) is actually spatially diffeomorphism invariant! Now in QGR the argument of the step function can be given a meaning as an operator (valued distribution) as we will see in the next section and the theta function of an operator can be defined through the spectral theorem. Since the spectrum of the θ-function consists only of $\{0, 1\}$, the spectrum of (82) should actually *coincide* with that of $\widehat{\text{Vol}}(R)$ [47]. A similar argument should also be valid with respect to Dirac observables commuting with the Hamiltonian constraint.

E) The existence of the area gap is also at the heart of the finiteness of the Bekenstein–Hawking entropy of black holes as we will see.

3 Selected Areas of Current Research

3.1 Quantum Dynamics

The Hamiltonian constraint C of QGR is, arguably, *the holy grail* of this approach to quantum gravity, therefore we will devote a substantial amount of space to this subject. In fact, unless one can quantize the Hamiltonian constraint, literally no further progress can be made so that it is important to know what its status is. From the explicit, non-polynomial expression (21) it is clear that a well-defined operator version of this object will be extremely hard to obtain and in fact this had been the major obstacle in the whole approach until the mid 90's. In particular, within the original ADM formulation only formal results were available. However, since with the new connection formulation also the non-polynomial kinematical operators of the previous section could be constructed, chances might be better.

At this point we include a brief account of the historical development of the subject in order to avoid confusion as one looks at older papers:
Originally one chose the Immirzi parameter as $\beta = \pm i$ and considered $\tilde{C} = \sqrt{\det(q)}$ rather than C because then $\tilde{C}$ is *actually a simple polynomial of only fourth order* (the "More" term disappears). Polynomiality was considered as mandatory. There were three problems with this idea:

1) The non-polynomiality was shifted from C into the reality conditions $A+\bar{A} = 2\Gamma(E)$ where the spin connection Γ is now a highly non-polynomial function of E. The operator version of this equation should be very hard to implement.
2) If A is complex, then we are dealing with an $SL(2,\mathbb{C})$ bundle rather than an $SU(2)$ bundle. Since $SL(2,\mathbb{C})$ is not compact, the mathematical apparatus of Sect. 2 is blown away.
3) Even formal trials to quantize $\tilde{C}$ resulted in either divergent, or background dependent operators.

In [27] it was suggested to keep β real which solves problems 1) and 2), however, then $\tilde{C}$ becomes even more complicated and anyway problem 3) is not cured. Finally in [48] it was shown that non-polynomiality is not necessarily an obstacle, even better, it is actually *required* in order to arrive at a well-defined operator: It was established that the reason for problem 3) is that $\tilde{C}$ is a scalar density of weight *two* while it was shown that only density weight *one* scalars have a chance to be quantized rigorously and background independently. Therefore the currently accepted point of view is that β should be real and that one uses the original unrescaled C rather than $\tilde{C}$.

A Possible New Mechanism for Avoiding UV Singularities in Background Independent Quantum Field Theories. Before we go into more details concerning [48], let us give a heuristic explanation *just why it happens that QGR may cure UV problems of QFT*, making the connection with the issue of the density weight just mentioned. Consider classical Einstein–Maxwell theory on $M = \mathbb{R} \times \sigma$ in its canonical formulation, then the Hamiltonian constraint gains an extra matter piece given for unit lapse $N = 1$ by

$$H_{EM} = \frac{1}{2e^2} \int_\sigma d^3x \; \overbrace{\frac{q_{ab}}{\sqrt{\det(q)}}}^{\text{Density weight -1}} \; \underbrace{[E^a E^b + B^a B^b]}_{\text{Density weight +2}} \quad . \tag{83}$$

Exercise 17.
Starting from the Lagrangean

$$L = -\frac{1}{4e^2} \sqrt{|\det(g)|} F_{\mu\nu} F_{\rho\sigma} g^{\mu\rho} g^{\nu\sigma} \,, \tag{84}$$

where $F = 2dA$ is the spacetime curvature of the Maxwell connection A with unit cm^{-1} and e is the electric charge in units such that $\alpha = \hbar e^2$ is the dimensionless fine-structure constant, perform the Legendre transform. With the electric field E^a being the momentum conjugate to the spatial piece A_a of A verify that the "Hamiltonian" is given by $-A_0 G + N^a V'_a + NC'$ where $G = \partial_a E^a$ is the Gauss law, $V'_a = F_{ab}E^b$ and C' is the integrand of (83) with $B^a = \epsilon^{abc}F_{bc}/2$ the magnetic field. Check that G' generates $U(1)$ gauge transformations while V'_a generates spatial diffeomorphisms where A_a, E^a transform as a one-form and a vector density of weight one respectively. Confirm that also B^a is a vector density of weight one.

As the exercise reveals, the geometry factor in (83) is a symmetric covariant tensor of rank two of density weight -1 due to the factor $\sqrt{\det(q)}$ in the denominator while the matter part is a symmetric contravariant tensor of rank two of density weight $+2$. That the resulting scalar has net density weight is $+1$ is no coincidence but a direct consequence of the diffeomorphism invariance or background independence of any matter theory coupled to gravity: only the integral over σ of a scalar density of weight $+1$ is spatially diffeomorphism invariant.

We can now quantize (83) in two ways:

1) In the first version we notice that if $g = \eta$ is the Minkowski metric, that is, $q_{ab} = \delta_{ab}$ then (83) reduces to the ordinary Maxwell Hamiltonian on Minkowski space. Thus we apply the formalism of QFT on a background spacetime, in this case Minkowski space, because we have fixed q_{ab} to the non-dynamical $\mathbb{C}$-number field δ_{ab} which is not quantized at all.
2) In the second version we keep q_{ab} dynamical and quantize it as well. Thus we apply QGR, a background independent quantization. Now q_{ab} becomes a field operator $\hat{q}_{ab}$ and the statement that the metric is flat can at most have a semiclassical meaning, that is, the expectation value of $\hat{q}_{ab}$ in a gravitational state is close to δ_{ab}.

Let us now sketch how these two different quantizations are performed and exactly pin-point how it happens that the first quantization is divergent while the second is finite.

1) *QFT on a background spacetime*
As we have said, the metric $q_{ab} = \delta_{ab}$ is now no longer a dynamical entity but just becomes a complex number. What we get is the usual Maxwell Hamiltonian operator

$$\hat{H}_M = \frac{1}{2e^2}\int_\Sigma d^3x\ \delta_{ab}[\hat{E}^a\hat{E}^b + \hat{B}^a\hat{B}^b]\,. \tag{85}$$

Notice the crucial difference with (83): The net density weight of the operator valued distribution in the integral is now $+2$ rather than $+1$! By switching off the metric as a dynamical field we have done a severe crime to the operator, because the net density weight $+2$ will be remembered by the operator in any faithful representation of the canonical commutation relations and leads to the following problem: The only coordinate density of weight one that one can construct is a δ-distribution (and derivatives thereof), thus for instance the operator $\hat{E}^a(x)$ is usually represented as a functional derivative which one can rewrite formally as

$$\alpha\delta/\delta A_a(x) = \alpha \sum_{y\in\Sigma} \delta(x,y)\partial/\partial A_a(y)\,. \tag{86}$$

The right hand side of (86) is a sum over terms each of which consists of a well-defined operator $Y_a(y) = \partial/\partial A_a(y)$ and a distributional prefactor $\delta(x,y)$. It is for this reason that expressions of the form $\hat{E}^a(x)\hat{E}^b(x)$ cannot be well-defined since we get products of distributions supported at the same point x and which result in divergent expressions of the form $\sum_{y,z}\delta(x,y)\delta(x,z)Y_a(y)Y_b(z) = \sum_y \delta(x,y)^2 Y_a(y)Y_b(y)$. The density weight two is correctly encoded in the term $\delta(x,y)^2 = \delta(0,0)\delta(x,y)$ which, however, is meaningless.
These heuristic arguments can of course be made precise: (85) is quantized on the Fock space $\mathcal{H}_{Fock}$ and one obtains

$$\hat{H}_M = :\hat{H}_M: + \hbar\int_\Sigma \underbrace{[\sqrt{-\Delta_x}\delta(x,y)]_{x=y}}_{\text{UV Singularity}}\,. \tag{87}$$

Here the colons stand for normal ordering. The UV (or short distance) singularity is explicitly identified as the coincidence limit $x = y$ of the integrand in the normal ordering correction. Therefore $\hat{H}_M$ is ill-defined on $\mathcal{H}_{Fock}$. Notice that even if the integrand would be finite, the integral suffers from an IR (or large volume) singularity if σ is not compact which comes from the fact that we are dealing with an infinite number of degrees of freedom. This singularity is, in contrast to the UV singularity, physical since it captures the vacuum energy of the universe which is of course infinite if the volume is.

2) *QFT coupled to QGR*
This time we keep the metric as a dynamical variable and quantize it. Thus instead of (85) we obtain something of the form

$$\hat{H}_{EM} = \frac{1}{2e^2}\int_\Sigma d^3x\, \widehat{\frac{q_{ab}}{\sqrt{\det(q)}}}\,[\hat{E}^a\hat{E}^b + \hat{B}^a\hat{B}^b]\,. \tag{88}$$

This time the net density weight is still +1. Now while the expression (86) is still valid and implies that there will be a product of δ-distributions in the numerator coming from the matter operator valued distributions, there is also a δ-distribution *in the denominator* due to the factor $\sqrt{\det(q)}$ which comes about as follows: As we already mentioned in Sect. 2.2 the volume functional in (81) admits a well-defined quantization of the form

$$\widehat{\mathrm{Vol}}(R)T_s = \ell_{\mathrm{P}}^3 \sum_{v\in V(\gamma(s))\cap R} \hat{V}_v T_s\,, \tag{89}$$

where $\hat{V}_v$ is a well-defined, dimensionless operator (not an operator valued distribution!) built from the vector fields v^j_S. Since $\text{Vol}(R)$ is the integral over R of $\sqrt{\det(q)}$ we conclude that $\sqrt{\det(q)}$ admits a quantization as an operator valued distribution, namely

$$\widehat{\sqrt{\det(q)}}(x)T_s = \ell_{\rm P}^3 \sum_{v\in V(\gamma(s))} \delta(x,v)\hat{V}_v T_s. \tag{90}$$

Now certainly (88) cannot be quantized on the Hilbert space $\mathcal{H}_{\rm kin}\otimes\mathcal{H}_{Fock}$ because $\mathcal{H}_{Fock}$ depends on a background metric (for instance through the Laplacian Δ) which is not available to us. However, we may construct a background independent Hilbert space $\mathcal{H}'_{\rm kin}$ for Maxwell theory which is completely identical to our $\mathcal{H}_{\rm kin}$, just that $SU(2)$ is replaced by $U(1)$ [48]. In $\mathcal{H}'_{\rm kin}$ the role of spin network states is played by charge network (CNW) states, that is, edges e are labelled by integers n_e (irreducibles of $U(1)$. Let us denote CNW's by $c=(\gamma,\boldsymbol{n}=\{n_e\}_{e\in E(\gamma)})$ and CNWF's by T'_c. Then a basis for the Einstein-Maxwell theory kinematical Hilbert space $\mathcal{H}_{\rm kin}\otimes\mathcal{H}'_{\rm kin}$ is given by the states $T_s\otimes T'_c$.
Now something very beautiful happens, which is not put in by hand but rather is a derived result: A priori the states T_s, T'_c may live on different graphs, however, unless the graphs are identical, the operator automatically (88) annihilates $T_s\otimes T'_c$ [49]. This is the mathematical manifestation of the following deep physical statement: *Matter can only exist where geometry is excited*. Indeed, if we have a gravitational state which has no excitations in a coordinate region R then the volume of that region as measured by the volume operator is identically zero. However, if a coordinate region has zero volume, then it is *physically simply not there, it is empty space*. Summarizing, the operator (88) is non-trivial only if $\gamma(s)=\gamma(c)$.
With this being understood, let us then sketch the action of (88) on our basis. One finds heuristically

$$\begin{aligned}\hat{H}_{EM}T_s\otimes T'_c = m_{\rm P}\sum_{v\in V(\gamma)}\ \sum_{e,e'\in E(\gamma), e\cap e'=v} \times \qquad\qquad & \qquad (91)\\ \times\int_\Sigma d^3x\Big[\frac{\hat{q}_{e,e'}}{\hat{V}_v}\,\underbrace{\frac{1}{\delta(x,v)}}_{\uparrow}\,T_s\Big]\otimes\Big[\underbrace{\delta(x,v)}_{\uparrow}\,\delta(y,v)Y^eY^{e'}T'_c]_{x=y}\Big], &\\ \text{Cancellation} &\end{aligned}$$

where $m_{\rm P}=\sqrt{\hbar/\kappa}$ is the Planck mass. Here $\hat{q}_{e,e'}$ and Y^e are well-defined, dimensionless operators (not distribution valued!) on $\mathcal{H}_{\rm kin}$ and $\mathcal{H}'_{\rm kin}$ respectively built from the right invariant vector fields R^j_e, R_e that enter the definition of the flux operators as in (49) and its analog for $U(1)$. The product of δ-distributions in the numerator of (91) has its origin again in the fact that the matter operator has density weight $+2$ certainly also in this representation and therefore has to be there, so nothing is swept under the rug! The

δ-distribution in the denominator comes from (90) and correctly accounts for the fact that the geometry operator has density weight -1. Again we have a coincidence limit $x = y$ which comes from a point splitting regularization and which in the background dependent quantization gave rise to the UV singularity. Now we see what happens: One of the δ-distributions in the numerator gets precisely cancelled by the one in the denominator leaving us with only one δ-distribution correctly accounting for the fact that the net density weight is $+1$. The integrand is then well-defined and the integral can be performed resulting in the *finite* expression

$$\hat{H}_{EM} T_s \otimes T'_c = \sum_{v\in V(\gamma)} \sum_{e,e'\in E(\gamma), e\cap e'=v} \times \left[\frac{\hat{q}_{e,e'}}{\hat{V}_v} T_s\right] \otimes [Y^e Y^{e'} T'_c]_{x=y} \,. \tag{92}$$

Notice that *finite* here means *non-perturbatively finite*, that is, not only finite order by order in perturbation theory (notice that in coupling gravity we have a highly interacting theory in front of us). Thus, comparing our non-perturbative result to perturbation theory the result obtained is comparable to showing that the perturbation series *converges*! Notice also that for non-compact σ the expression (92) possibly has the physically correct IR divergence coming from a sum over an infinite number of vertices.

EXERCISE 18.
Recall the Fock space quantization of the Maxwell field and verify (87).

This ends our heuristic discussion about the origin of UV finiteness in QGR. The crucial point is obviously the density weight of the operator in question which should be precisely $+1$ in order to arrive at a well-defined, background independent result: Higher density weight obviously leads to more and more divergent expressions, lower density weight ends in zero operators.

Sketch of a Possible Quantization of the Hamiltonian Constraint. We now understand intuitively why the rescaled Hamiltonian constraint $\tilde{C}$ had no chance to be well-defined in the quantum theory: It is similar to (87) due to its density weight $+2$. The same factor $1/\sqrt{\det(q)}$ that was responsible for making (88) finite also makes the original, non-polynomial, unrescaled Hamiltonian constraint $C = \tilde{C}/\sqrt{\det(q)}$ finite. We will now proceed to some details how this is done, avoiding intermediate divergent expressions such as in (91).

The essential steps can already be explained for the first term in (21) so let us drop the “More” term and consider only the integrated first term

$$C_E(N) = \frac{1}{\kappa} \int_{\Sigma} d^3x N \frac{F^j_{ab} \epsilon_{jkl} E^a_k E^b_l}{\sqrt{|\det(E)|}} \,. \tag{93}$$

Let us introduce a map

$$R: \ \Sigma \to \mathcal{O}(\sigma); \ x \mapsto R_x \tag{94}$$

where $\mathcal{O}(\Sigma)$ denotes the set of open, compactly supported, connected and simply connected subsets of Σ and $R_x \in \mathcal{O}(\Sigma)$ is constrained by the requirement that $x \in R_x$. We define the volumes of the R_x by

$$V(x) := \mathrm{Vol}(R_x) = \int_{R_x} d^3y \sqrt{|\det(E)|}(y)\,. \tag{95}$$

Then, up to a numerical prefactor we may write (93) in the language of differential forms and in terms of a Poisson bracket as

$$C_E(N) = \frac{1}{\kappa^2} \int_\Sigma N \mathrm{Tr}(F \wedge \{A, V\})\,. \tag{96}$$

Exercise 19.
Verify that (95) is really the volume of R_x and (96).

The reasoning behind (96) was to move the factor $1/\sqrt{\det(q)}(x)$ from the denominator into the numerator by using a Poisson bracket. This will avoid the δ-distribution in the denominator as in (91) and has the additional advantage that $\sqrt{\det(q)}$ now appears smeared over R_x so that one obtains an operator, not a distribution. Thus, the idea is now to replace the function $V(x)$ by the well-defined operator $\widehat{\mathrm{Vol}}(R_x)$ and the Poisson bracket by a commutator divided by $i\hbar$. The only thing that prevents us from doing this is that the operators A_a^j, F_{ab}^j do not exist on $\mathcal{H}_{\mathrm{kin}}$. However, they can be regularized in terms of holonomies as follows:
Given tangent vectors $u, v \in T_x(\sigma)$ we define one parameter homotopies of paths and loops of triangle topology

$$\epsilon \mapsto p^u_{\epsilon,x},\ \alpha^{uv}_{\epsilon,x} \tag{97}$$

respectively with $b(p^u_{\epsilon,x}) = b(\alpha^{uv}_{\epsilon,x}) = x$ and $(\dot{p}^u_{\epsilon,x})_x = (\dot{\alpha}^{uv}_{\epsilon,x})_{x+} = \epsilon u,\ (\dot{\alpha}^{uv}_{\epsilon,x})_{x-} = -\epsilon v$ (left and right derivatives at x). Then for *smooth* connections $A \in \mathcal{A}$ the Ambrose–Singer theorem tells us that

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon}[A(p^u_{\epsilon,x}) - 1] = u^a A_a(x), \quad \lim_{\epsilon\to 0} \frac{1}{\epsilon^2}[A(\alpha^{uv}_{\epsilon,x}) - 1] = \frac{1}{2} u^a v^b F_{ab}(x)\,. \tag{98}$$

Exercise 20.
Verify (98) by elementary means, using directly the differential equation (38).
Hint:
For sufficiently small ϵ we have up to ϵ^2 corrections $p^u_{\epsilon,x}(t) = x + \epsilon t u$ and

$$\alpha^{uv}_{\epsilon,x}(t) = x + \epsilon \begin{cases} tu & ,\quad t \in [0,1] \\ u/3 + (t-1)(v-u) & ,\quad t \in [1,2] \\ (3-t)v & ,\quad t \in [2,3]. \end{cases}$$

Thus, given a triangulation τ_ϵ of σ, that is, a decomposition of σ into tetrahedra Δ with base points $v(\Delta)$, edges $p_I(\Delta),\ I = 1,2,3$ of Δ of the type $p^u_{\epsilon,v}$ starting at v and triangular loops $\alpha_{IJ}(\Delta) = p_I(\Delta) \circ a_{IJ}(\Delta) p_J(\Delta)^{-1}$ of the type $\alpha^{uv}_{\epsilon,v}$

where the *arcs* $a_{IJ}(\Delta)$ comprise the remaining three edges of Δ, it is easy to show, using (98), that up to a numerical factor

$$C_E^{\tau_\epsilon}(N) = \frac{1}{\kappa^2} \sum_{\Delta \in \tau_\epsilon} N(v(\Delta)) \sum_{IJK} \epsilon^{IJK} \mathrm{Tr}(A(\alpha_{IJ}(\Delta)A(p_K(\Delta))\{A(p_K(\Delta))^{-1}, V(v)\} \tag{99}$$

tends to $C_E(N)$ as $\epsilon \to 0$ (in this limit the triangulation gets finer and finer).

EXERCISE 21.
Verify this statement.

The expression (99) can now be readily quantized on $\mathcal{H}_{\mathrm{kin}}$ because holonomies and volume functionals are well-defined operators. However, we must remove the regulator ϵ in order to arrive at a quantization of (96). Now the regulator can be removed in many inequivalent ways because there is no unique way to refine a triangulation. Moreover, we must specify in which operator topology $\hat{C}^{\tau_\epsilon}(N)$ converges. The discussion of these issues is very complicated and the interested reader is referred to [48] for the detailed arguments that lead to the following solution:

i) *Triangulation*
First of all we define the operator explicitly on the SNW basis T_s. In order for the refinement limit to be non-trivial, it turns out that the triangulation must be refined in such a way that $\gamma(s) \subset \tau_\epsilon$ for sufficiently small ϵ. This happens essentially due to the volume operator which has non-trivial action only at vertices of graphs. Thus the refinement must be chosen depending on s. This is justified because classically all refinements lead to the same limit. One might worry that this does not lead to a linear operator, however, this is not the case because it is defined on a basis.

ii) *Operator Topology*
The limit $\epsilon \to 0$ exists in the following sense:
Let $\mathcal{D}^*_{\mathrm{Diff}} \subset \mathcal{D}^*_{\mathrm{kin}}$ be the space of solutions of the diffeomorphism constraint. We say that a family of operators $\hat{O}_\epsilon$ converges to an operator $\hat{O}$ on $\mathcal{H}_{\mathrm{kin}}$ in the *uniform-weak-Diff*-topology* provided that for each $\delta > 0$ and for each $l \in \mathcal{D}^*_{\mathrm{Diff}}$, $f \in \mathcal{D}_{\mathrm{kin}}$ there exists $\epsilon(\delta) > 0$ *independent of* l, f such that

$$|l([\hat{O}_\epsilon - \hat{O}]f)| < \delta \;\; \forall \epsilon < \epsilon(\delta)\,. \tag{100}$$

This topology is of course motivated by physical considerations: Since the operator is unbounded, the uniform (i.e. operator norm) topology is too strong. The strong or weak topologies (pointwise convergence in Hilbert space norm or as matrix elements) give a trivial (zero) limit (exercise!). Thus one is naturally led to * topologies. The maximal dual space on which to build a topology would be $\mathcal{D}^*_{\mathrm{kin}}$ but one can check that the limit does not exist even pointwise in $\mathcal{D}^*_{\mathrm{kin}}$. Thus one is looking for suitable subspaces thereof. The natural, physically motivated choice is, of course, the space $\mathcal{D}^*_{\mathrm{Diff}}$

which is singled out by the spatial diffeomorphism constraint. The reason for why have required uniform convergence in (100) is that this excludes the existence of the limit for larger spaces $\mathcal{D}^*_{\rm Diff} \subset \mathcal{D}^*_\star \subset \mathcal{D}^*_{\rm kin}$.

The end result is

$$\hat{C}^\dagger_E(N) = m_{\rm P} \sum_{v \in V(\gamma(s))} N(v) \sum_{\substack{e,e',e'' \in E(\gamma(s)) \\ e \cap e' \cap e'' = \{v\}}} \times$$
$$\times \{{\rm Tr}([A(\alpha_{\gamma(s),v,e,e'}) - (A(\alpha_{\gamma(s),v,e,e'}))^{-1}]A(p_{\gamma(s),v,e''})[A(p_{\gamma(s),v,e''})^{-1}, \hat{V}_v])$$
$$+\text{cyclic permutation in}\{e,e',e''\}\}\ T_s\,. \tag{101}$$

The meaning of the loops $\alpha_{\gamma,v,e,e'}$ and paths $p_{\gamma,v,e''}$ that appear in this sum over vertices and triples of edges incident at them is best explained in the following Fig. 11. Their precise specification makes use of the axiom of choice and is diffeomorphism covariant, that is, for $\varphi \in {\rm Diff}^\omega(\Sigma)$, e.g. the loops $\alpha_{\gamma,v,e,e'}$ and $\alpha_{\varphi(\gamma),\varphi(v),\varphi(e),\varphi(e')}$ are analytically diffeomorphic. Moreover, the *arcs* $a_{\gamma,v,e,e'}$ defined by

$$\alpha_{\gamma,v,e,e'} = p_{\gamma,v,e} \circ a_{\gamma,v,e,e'} \circ p^{-1}_{\gamma,v,e'} \tag{102}$$

are such that also $\gamma \cup a_{\gamma,v,e,e'}$ and $\varphi(\gamma) \cup a_{\varphi(\gamma),\varphi(v),\varphi(e),\varphi(e')}$ are analytically diffeomorphic. The adjoint in (101) is due to the fact that $C_E(N)$ is classically

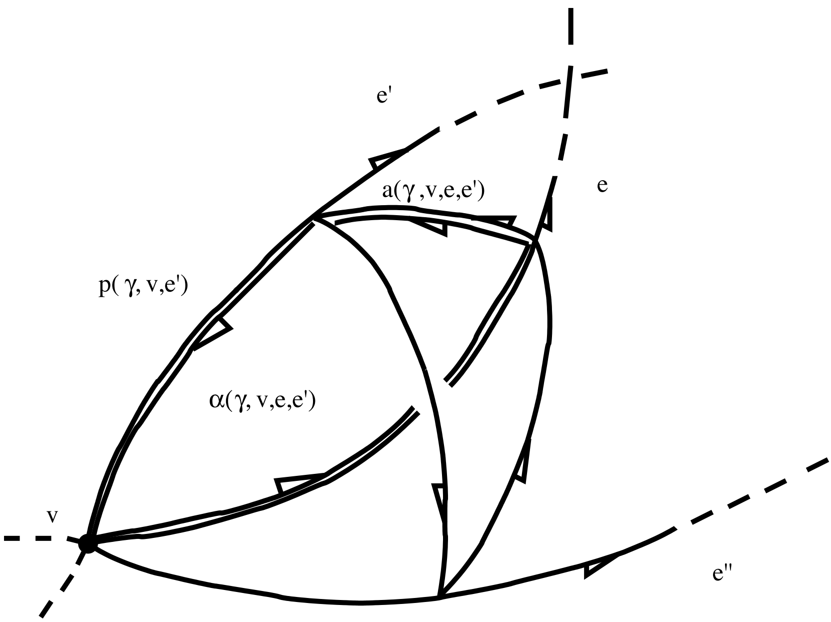


Fig. 11. Meaning of the loop, path and arc assignment of the Hamiltonian constraint. Notice how a tetrahedron emerges from those objects, making the link with the triangulation. The broken lines indicate possible other edges or continuations thereof

real-valued, so we are quantizing $\overline{C_E(N)}$ as well. The operator (101) is not symmetric, however, its adjoint is densely defined on $\mathcal{D}_{\rm kin}$ and it is therefore closable. Usually one requires real valued functions to become self-adjoint operators because then by the spectral theorem the spectrum (possible measurement values) is a subset of the real line. However, this argument is void when we are only interested in the kernel of the operator ("zero eigenvalue").

Exercise 22.
Verify that $\hat{C}^\dagger_E(N)$ is not symmetric but it is, together with $\hat{C}_E(N)$, densely defined on $\mathcal{D}_{\rm kin}$. Show that if real valued constraints C_I form a Poisson algebra $\{C_I, C_J\} = f_{IJ}\ ^K C_K$ with non-trivial, real valued structure functions such that $\{f_{IJ}\ ^K, C_K\}_{\{C_L=0\}} \neq 0$, then $\hat{C}_I, \hat{f}_{IJ}\ ^K$ must not be both symmetric in order for the quantum algebra to be free of anomalies. Conclude that the failure of (101) to be symmetric is likely to be required for reasons of consistency.

The fact the loop $\alpha_{\gamma,v,e,e'}$ is not shrunk to v as one would expect is of course due to our definition of convergence, in fact, an arbitrary loop assignment (γ, v, e, e') $\mapsto \alpha_{\gamma,v,e,e'}$ that has the same diffeomorphism invariant characteristics is allowed, again because in a diffeomorphism invariant theory there is no notion of "closeness" of $\alpha_{\gamma,v,e,e'}$ to v. Notice that the operator $\hat{C}_E(N)$ is *defined on* $\mathcal{H}_{\rm kin}$ using the axiom of choice and not on diffeomorphism invariant states as it is sometimes misleadingly stated in the literature [50]. In fact, it cannot be because the dual operator $\hat{C}'_E(N)$ defined by

$$[\hat{C}'_E(N)l](f) := l(\hat{C}^\dagger_E(N)f) \tag{103}$$

for all $f \in \mathcal{D}_{\rm kin}$, $l \in \mathcal{D}^*_{\rm kin}$ does not preserve $\mathcal{D}^*_{\rm Diff}$ as is expected from the classical Poisson algebra $\{V, C\} \propto C \neq V$. If one wants to take this dual point of view then one is forced to introduce a larger space $\mathcal{D}^*_\star$ which is preserved but which does not solve the diffeomorphism constraint and is therefore unphysical. This has unnecessarily given rise to a large amount of confusion in the literature and should be abandoned.

As we have said, the loop assignment is to a very large extent arbitrary at the level of $\mathcal{H}_{\rm kin}$ and represents a serious *quantization ambiguity*, it cannot even be specified precisely because we are using the axiom of choice. However, at the level of $\mathcal{H}_{\rm phys}$ this ambiguity evaporates to a large extent because all choices that are related by a diffeomorphism result in the same solution space *to all constraints* defined by elements $l \in \mathcal{D}^*_{\rm Diff}$ which satisfy in addition

$$[\hat{C}'_E(N)l](f) = l(\hat{C}^\dagger_E(N)f) = 0 \ \forall \ N \in C^\infty_0(\Sigma),\ f \in \mathcal{D}_{\rm kin}\,, \tag{104}$$

where $C^\infty_0(\Sigma)$ denotes the smooth functions of compact support. Thus the solution space $\mathcal{D}^*_{\rm phys}$ will depend only on the spatially diffeomorphism invariant characteristics of the loop assignment *which can be specified precisely* [48], it essentially characterizes the amount by which the arcs knot the original edges of the graph. Besides this remaining ambiguity there are also factor ordering ambiguities *but no singularities* some of which are discussed in [51].

Let us list without proof some of the properties of this operator:

i) *Matter Coupling*
Similar Techniques can be applied to the case of (possibly supersymmetric) matter coupled to GR [48].

ii) *Anomaly-Freeness*
The constraint algebra of the Hamiltonian constraint with the spatial diffeomorphism constraint and among each other is *mathematically consistent*. From the classical constraint algebra $\{V, C\} \propto C$ we expect that $\hat{V}(\varphi)\hat{C}_E^\dagger(N)\hat{V}(\varphi)^{-1} = \hat{C}_E(\varphi^* N)$ for all diffeomorphisms φ. However, this is just the statement of the loop assignment being diffeomorphism covariant which can be achieved by making use of the axiom of choice. Next, from $\{C, C\} \propto V$ we expect that the dual of $[\hat{C}_E^\dagger(N), \hat{C}_E^\dagger(N')] = [\hat{C}_E(N'), \hat{C}_E(N)]^\dagger$ annihilates the elements of $\mathcal{D}^*_{\text{Diff}}$. This can be explicitly verified [48]. We stress that $[\hat{C}_E^\dagger(N), \hat{C}_E^\dagger(N')]$ is *not zero*, the algebra of Hamiltonian constraints is *not Abelean* as it is sometimes misleadingly stated in the literature. The commutator is in fact explicitly proportional to a diffeomorphism.

iii) *Physical States*
There is a rich space of rigorous solutions to (104) and a precise algorithm for their construction has been developed [48].

iv) *Intuitive Picture*
The Hamiltonian constraint acts by annihilating and creating spin degrees of freedom and therefore the dynamical theory obtained could be called “Quantum Spin Dynamics (QSD)” in analogy to “Quantum Chromodynamics (QCD)” in which the Hamiltonian acts by creating and annihilating colour degrees of freedom. In fact we could draw a crude analogy to Fock space terminology as follows: The (perturbative) excitations of QCD carry a continuous label, the mode number $k \in \mathbb{R}^3$ and a discrete label, the occupation number $n \in \mathbb{N}$ (and others). In QSD the continuous labels are the edges e and the discrete ones are spins j (and others). So we have something like a non-linear Fock representation in front of us.
Next, when solving the Hamiltonian constraint, that is, when *integrating the Quantum Einstein Equations*, one realizes that one is not dealing with a (functional) partial differential equation but rather with a (functional) partial difference equation. Therefore, when understanding coordinate time as measured how for instance volumes change, we conclude that also time evolution is necessarily *discrete*. Such discrete time evolution steps driven by the Hamiltonian constraint assemble themselves into what nowadays is known as a spin foam. A spin foam is a four dimensional complex of two dimensional surfaces where each surface is to be thought of as the world sheet of an edge of a SNW and it carries the spin that the edge was carrying before it was evolved[7].
Another way of saying this is that a spin foam is a complex of two-surfaces labelled by spins and when cutting a spin foam with a spatial three-surface Σ

[7] Thus, a spin foam model can be thought of as a background independent string theory!

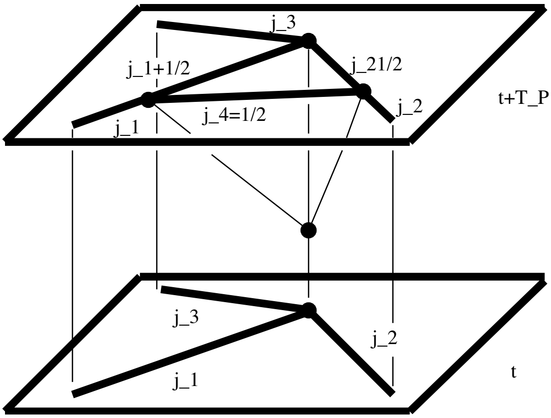


Fig. 12. Emergence of a spin foam from a SNW by the action of the Hamiltonian constraint

one obtains a SNW. If one uses two such surfaces $\Sigma_t, \Sigma_{t+T_{\mathrm{P}}}$ where $T_{\mathrm{P}} = \ell_{\mathrm{P}}/c$ is the Planck time then one rediscovers the discrete time evolution of the Hamiltonian constraint. These words are summarized in Fig. 12.

While these facts constitute a promising hint that the Hilbert space $\mathcal{H}_{\mathrm{kin}}$ could in fact *support the quantum dynamics of GR*, there are well-taken concerns about the physical correctness of the operator $\hat{C}_E^\dagger(N)$:

The problem is that one would like to see more than that the commutator of two dual Hamiltonian constraints annihilates diffeomorphism invariant states, one would like to see something of the kind

$$[\hat{C}_E^\dagger(N), \hat{C}_E^\dagger(N')] = i\ell_{\mathrm{P}}^2 [\widehat{\int_\Sigma d^3x [NN'_{,a} - N_{,a}N'] q^{ab} V_b}]. \tag{105}$$

The reason for this is that then one would be sure that $\hat{C}_E^\dagger(N)$ generates the correct quantum evolution. While this requirement is not necessary, it is certainly sufficient and would be reassuring[8]. There are two obstacles that prevent us from rewriting the left hand side of (105) in terms of the right hand side.

1) The one parameter groups $s \mapsto \hat{V}(\varphi_s^u)$ of unitarities where φ_s^u are the one parameter groups of diffeomorphisms defined by the integral curves of a vector field u are not weakly continuous, therefore a self-adjoint generator $\hat{V}(u)$ that we would like to see on the right hand side of (105) simply does not exist.

[8] Example: Suppose that C_a are the angular momentum components for a particle in in $\mathbb{R}^3$ with classical constraint algebra $\{C_a, C_b\} = \epsilon_{abc}C_c$. Introduce polar coordinates and define the non-self adjoint operators $\hat{C}_1 = i\hbar\partial/\partial\theta$, $\hat{C}_2 = i\hbar\partial/\partial\phi$, $\hat{C}_3 = 0$. Then the quantum constraint algebra is Abelean and does not at all resemble the classical one, however, the physical states are certainly the correct ones, functions that depend only on the radial coordinate.

Exercise 23.
Recall Stone's theorem about the existence of the self-adjoint generators of weakly continuous one-parameter unitary groups and verify that $\hat{V}(\varphi_s^u)$ is not weakly continuous on $\mathcal{H}_{\mathrm{kin}}$.

2) One can quantize the right hand side of (105) by independent means and it does annihilate $\mathcal{D}^*_{\mathrm{Diff}}$ [48], however, that operator does not resemble the left hand side in any obvious way. The reason for this is that even classically it takes a A4 page of calculation in order to rewrite the Poisson bracket $\{C_E(N), C_E(N')\}$ as in (105) with V_a given by (21). The manipulations that must be performed in order to massage the Poisson bracket into the desired form involve a) integrations by part, b) writing F_{ab} in terms of A_a, c) derivatives of $\sqrt{\det(q)}$, d) multiplying fractions by functions in both numerator and denominator, e) symmetry arguments in order to see that certain terms cancel etc. (exercise!). These steps are obviously difficult to perform with operators.

In summary, there is no mathematical inconsistency, however, there are doubts about the physical correctness of the Hamiltonian constraint operator presently proposed although no proof exists so far that it is necessarily wrong. In order to make progress on this issue, it seems that we need to develop first a semiclassical calculus for the theory, more precisely, we need coherent states so that expectation values of operators and their commutators can be replaced, up to $\hbar$ corrections, by their classical values and Poisson brackets respectively for which then the manipulations listed in 2) above can be carried out. If that is possible, and the outcome of these calculations is the expected one, possibly after changing the operator by making use of the available quantization ambiguities, then one would be able to claim that one has indeed constructed a quantum theory of GR with the correct classical limit. Only then can one proceed to solve the theory, that is, to construct solutions, the physical inner product and the Dirac observables. The development of a semiclassical calculus is therefore one of the "hot" research topics at the moment.

Another way to get confidence in the quantization method applied to the Hamiltonian constraint is to study model systems for which the answer is known. This has been done for 2+1 gravity [48] and for quantum cosmology to which we turn in the next section.

3.2 Loop Quantum Cosmology

A New Approach To Quantum Cosmology. The traditional approach to quantum cosmology consists in a so-called mini-superspace quantization, that is, one imposes certain spacetime Killing symmetries on the metric, plugs the symmetric metric into the Einstein Hilbert action and obtains an effective action which depends only on a finite number of degrees of freedom. Then one canonically quantizes this action. Thus one symmetrizes before quantization. These models are of constant interest and have natural connections to inflation. See e.g. [52] for recent reviews.

What is not perfect about these models is that 1) not only do they switch off all but an infinite number of degrees of freedom, but 2) also the quantization method applied to the reduced model usually is quite independent from that applied to the full theory. A fundamental approach to quantum cosmology will be within the full theory and presumably involves the construction of semiclassical physical states whose probability amplitude is concentrated on, say a Friedmann–Robertson–Walker (FRW) universe. This would cure both drawbacks 1) and 2). At the moment we cannot really carry out such a programme since the construction of the full theory is not yet complete. However, one can take a more modest, hybrid approach, where while dealing only with a finite number of degrees of freedom one *takes over all the quantization machinery from the full theory!* Roughly speaking, one works on the space $\mathcal{H}_{\rm kin}$ of the full theory but considers only states therein which satisfy the Killing symmetry. Hence one symmetrizes after quantization which amounts to considering only a finite subset of holonomies and fluxes. This has the advantage of leading to a solvable model while preserving pivotal structures of the full theory, e.g. the volume operator applied to symmetric states will still have a discrete spectrum as in the full theory while in the traditional approaches it is continuous. Such a programme has been carried out in great detail by Bojowald in a remarkable series of papers [53] and his findings are indeed spectacular, should they extend to the full theory: It turns out that the details of the quantum theory are *drastically different* from the traditional minisuperspace approach. In what follows we will briefly describe some of these results, skipping many of the technical details.

Spectacular Results. Consider the FRW line element (in suitable coordinates)

$$ds^2 = -dt^2 + R(t)^2[\frac{dr^2}{1-kr^2} + r^2 d\Omega_2^2] =: -dt^2 + R(t)^2 q^0_{ab} dx^a dx^b. \qquad (106)$$

The universe is closed/flat/open for $k = 1/0/-1$. The only dynamical degree of freedom left is the so-called scale factor $R(t)$ which describes the size of the universe and its conjugate momentum. The classical big bang singularity corresponds to the fact that the Einstein equations predict that $\lim_{t\to 0} R(t) = 0$ at which the metric (106) becomes singular and the inverse scale factor $1/R(t)$ blows up (the curvature will be $\propto 1/R(t)^2$ so this singularity is a true curvature singularity).

We are interested in whether the curvature singularity $1/R \to \infty$ exists also in the quantum theory. To study this we notice that for (85) $\det(q) = R^6 \det(q^0)$. Hence, up to a numerical factor this question is equivalent to the question whether the operator corresponding to $1/\sqrt[6]{\det(q)}$, when applied to symmetric states, is singular or not. However, we saw in the previous section that one can trade a negative power of $\det(q)$ by a Poisson bracket with the volume operator. In [53] precisely this, for the Hamiltonian constraint, *essential quantization technique* is applied which is why *this model tests some aspects of the quantization of the Hamiltonian constraint.* Now it turns out that this operator, applied to symmetric states, leads to an operator $\widehat{\frac{1}{R}}$ which is diagonalized

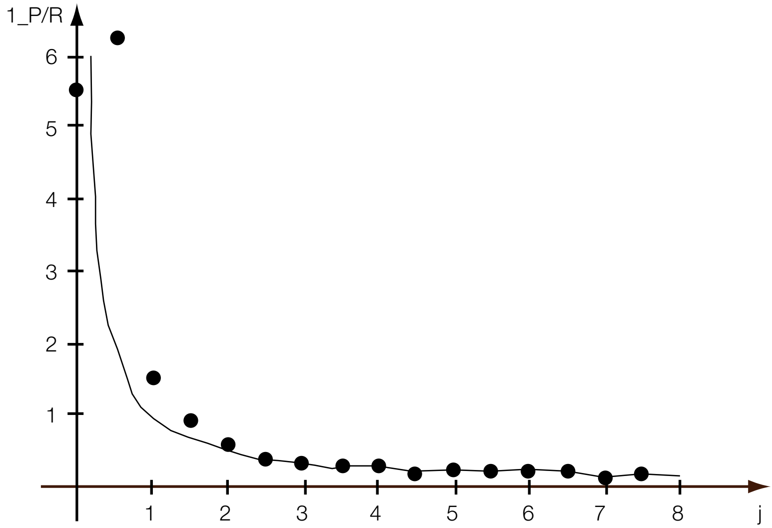


Fig. 13. Spectrum of the inverse scale factor

by (symmetric) SNWF's *and the spectrum is bounded*! In Fig. 13 we plot the qualitative behaviour of the eigenvalues $\ell_{\mathrm{P}}\lambda_j$ as a function of j where j is the spin label of a gauge invariant SNWF with a graph consisting of one loop only (that only such states are left follows from a systematic analysis which defines what a symmetric SNWF is). One can also quantize the operator $\hat{R}$ and one sees that its eigenvalues are essentially given by $j\ell_{\mathrm{P}}$ up to a numerical factor. Thus the classical singularity corresponds to $j = 0$ and one expects the points $\lambda_j \ell_p$ at the values $1/j$ on the curve ℓ_{P}/R. Evidently the spectrum is discrete (pure point) and bounded, at the classical singularity it is *finite*. In other words, the quantum universe never decreases to zero size. For larger j, in fact already for R of the order of ten Planck lengths and above, the spectrum follows the classical curve rather closely hinting at a well-behaved classical limit (correspondence principle).

Even more is true: One can in fact quantize the Hamiltonian constraint by the methods of the previous section and solve it exactly. One obtains an eighth order difference equation (in j). The solution therefore depends non-trivially on the initial condition. What is surprising, however, is the fact that only one set of initial conditions leads to the correct classical limit, thus in loop quantum cosmology initial conditions are *derived rather than guessed*. One can even propagate the quantum Einstein equations through the classical singularity and arrives at the picture of a bouncing universe.

Finally one may wonder whether these results are qualitatively affected by the operator ordering ambiguities of the Hamiltonian constraint. First of all one finds that these results hold only if one orders the loop in (101) to the left of the volume operator as written there. However, one is not forced to work with the holonomy around that loop in the fundamental representation of $SU(2)$, there is some flexibility [51] and one can choose a different one, say j_0. It turns out that the value j_0 influences the onset of classical behaviour, that is, the higher j_0 the higher the value $j(j_0)$ from which on the spectrum in Fig. 13 lies on the curve $1/j$. Now this is important when one couples, say scalar matter because the operator $\widehat{\frac{1}{R}}$ enters the matter part of the Hamiltonian constraint and modifies the resulting effective equation for $R(t)$ in the very early phase of the universe and leads to a *quantum gravity driven inflationary period* whose duration gets larger with larger j_0!

Thus, loop quantum cosmology not only confirms aspects of the quantization of the Hamiltonian constraint but also predicts astonishing deviations from standard quantum cosmology which one should rederive in the full theory.

3.3 Path Integral Formulation: Spin Foam Models

Spin Foams from the Canonical Theory. Spin Foam models are the fusion of ideas from topological quantum field theories and loop quantum gravity, see e.g. [54] for a review, especially the latest, most updated one by Perez. The idea that connects these theories is actually quite simple to explain at an heuristic level:

If we forget about 1) all functional analytic details, 2) the fact that the operator valued distributions corresponding to the Hamiltonian constraint $\hat{C}(x)$ do not mutually commute for different $x \in \sigma$ and 3) that the Hamiltonian constraint operators $\hat{C}(N)$ are certainly not self-adjoint, at least as presently formulated, then we can formally write down the complete space of solutions to the Hamiltonian constraint as a so-called "rigging map" (see e.g. [1])

$$\bar{\eta} : \mathcal{D}_{\text{kin}} \to \mathcal{D}^*_{\text{phys}};\ f \mapsto \delta[\hat{C}]\ f := [\prod_{x\in\Sigma} \delta(\hat{C}(x))\ f] \tag{107}$$

(where $\bar{\eta} = \text{c.c} \cdot \eta$ is the complex conjugate of the actual anti-linear rigging map). Here the δ-distribution of an operator is defined via the spectral theorem (assuming the operator to be self-adjoint). Notice that we do not need to order the points $x \in \sigma$ as we assumed the $\hat{C}(x)$ to be mutually commuting for the moment and only under this assumption it is true that, at least formally $\bar{\eta}[f](\hat{C}(N)f') = 0$ (exercise)[9]. Now we use the formula $\delta(x) = \int_{\mathbb{R}} \frac{dk}{2\pi} e^{ikx}$ to write the functional δ-distribution $\delta[\hat{C}]$ as a *path integral*

$$\delta[\hat{C}] = \int_{\mathcal{N}'} [DN] e^{i\hat{C}(N)}\,, \tag{108}$$

where we have neglected an infinite constant as usual in this formal business. Here $\mathcal{N}'$ is the space of lapse functions at a fixed time. Let us introduce also the

[9] At an even more formal level $\bar{\eta}[f]$ is also a solution in the non-commuting case if, as is the case with the currently proposed $\hat{C}$, the algebra with the spatial diffeomorphism constraint closes

space of lapses with arbitrary time dependence $\mathcal{N}_{t_1,t_2}$ in $t \in [t_1, t_2]$. Then, up to an infinite constant one can verify that

$$\delta[\hat{C}] = \int_{\mathcal{N}_{t_1}^{t_2}} [DN] e^{i \int_{t_1}^{t_2} dt \int_\Sigma d^3x N(x,t) \hat{C}(x)} . \tag{109}$$

The rigging map machinery then tells us that the scalar product on the image of the rigging map is simply given by

$$< \bar{\eta}(f), \bar{\eta}(f') >_{\mathrm{phys}} := < f, \bar{\eta}(f') >_{\mathrm{kin}} = \int_{\mathcal{N}_{t_1}^{t_2}} [DN] < f, e^{i \int_{t_1}^{t_2} dt \hat{C}(N_t)} f' >_{\mathrm{kin}} . \tag{110}$$

This formula looks like a propagator formula, that is, like a transition amplitude between an initial state f' on Σ_{t_1} and a final state f on Σ_{t_2} after a multi-fingered time evolution generated by $\hat{C}(N_t)$. In fact, if we use the Taylor expansion of the exponential function and somehow regularize the path integral then the expansion coefficients $< T_s, \hat{C}(N_t)^n T' s' >_{\mathrm{kin}}$ can be interpreted as probability amplitude of the evolution of the SNW state $T'_{s'}$ to reach the SNW state T_s after n time steps (recall Fig. 12).

Now by the usual formal manipulations that allow us to express a unitary operator $e^{i(t_2-t_1)\hat{H}}$ as a path integral over the classical pase space $\mathcal{M}$ (the rigorous version of which is the Feynman–Kac formula, e.g. [56]) one can rewrite (110) as

$$< \bar{\eta}(f), \bar{\eta}(f') >_{\mathrm{phys}} = \int [DN \; D\boldsymbol{N} \; D\Lambda \; DA \; DE] < f, e^{iS} f' >_{\mathrm{kin}} , \tag{111}$$

where S is the Einstein-Hilbert action written in canonical form in terms of the variables A, E, that is

$$S = \int_{\mathbb{R}} dt \int_\Sigma d^3x \{ \dot{A}_a^j E_j^a - [-\Lambda^j G_j + N^a V_a + NC] \} \tag{112}$$

and we have simultaneously included also projections on the space of solutions to the Gauss and vector constraint. Now the action (112) is the $3+1$ split of the following *covariant action*

$$S = \int_M \{ \Omega_{IJ} \wedge [\epsilon^{IJKL} - \beta^{-1} \eta^{IK} \eta^{JK}] e_K \wedge e_L \} \tag{113}$$

discovered in [57] where β is the Immirzi parameter. Here Ω_{IJ} is the (antisymmetric) curvature two-form of an (antisymmetric) $SL(2, \mathbb{C})$ connection one-form ω_{IJ} with Lorentz indices $I, J, K, .. = 0, 1, 2, 3$, η is the Minkowski metric and e^I is the co-tetrad one-form. The first term in (113) is called the Palatini action while the second term is topological (a total differential modulo the equations of motion). The relation between the four-dimensional fields ω_μ^{IJ}, e_μ^I (40 components) and the three-dimensional fields $A_a^j, E_j^a, \Lambda^j, N, N^a$ (25 components) can only be established if certain so-called second class constraints [20] are solved.

Spin Foams and BF-Theory. Thus, it is formally possible to write the inner product between physical states as a *covariant* path integral for the classical canonical action and using only the kinematical inner product, thus providing a bridge between the covariant and canonical formalism. However, this bridge is far from being rigorously established as we had to perform many formal, unjustified manipulations. Now rather than justifying the steps that lead from $\hat{C}$ to (111) one can turn the logic upside down and *start* from a manifestly covariant formulation and *derive* the canonical formulation. This is the attitude taken by people working actively on spin foam models. Thus, let us forget about the topological term in (113) and consider only the Palatini term. Then the Palatini action has precisely the form of a BF-action

$$S_{BF} = \int_M \Omega_{IJ} \wedge B^{IJ} \tag{114}$$

just that the (antisymmetric) two-form field B^{IJ} is not arbitrary (it would have 36 independent components), it has to come from a tetrad with only 16 independent components, that is, it has to be of the form $\epsilon^{IJKL} e_K \wedge e_L$.

Exercise 24.
Show that the condition that B comes from a tetrad is almost[10] equivalent to the *simplicity constraint*

$$\epsilon_{IJKL} B^{IJ}_{\mu\nu} B^{KL}_{\rho\sigma} = c\epsilon_{\mu\nu\rho\sigma} \tag{115}$$

for some spacetime scalar density c of weight one.

The reasoning is now as follows: BF-theory without the constraint (115) is a topological field theory, that is, it has no local degrees of freedom. Therefore quantum BF-theory is not really a QFT but actually a quantum mechanical system and can therefore be handled much more easily than gravity. Let us now write an action equivalent to the Palatini action given by

$$\begin{aligned} S'_P[\omega, B, \Phi] &= S_{BF}[\omega, B] + S_I[B, \Phi] \\ S_I[B, \Phi] &:= \int_M \Phi^{\mu\nu\rho\sigma} \epsilon_{IJKL} B^{IJ}_{\alpha\beta} B^{KL}_{\gamma\delta} [\delta^\alpha_\mu \delta^\beta_\nu \delta^\gamma_\rho \delta^\delta_\sigma - \frac{1}{4!} \epsilon^{\alpha\beta\gamma\delta} \epsilon_{\mu\nu\rho\sigma}] , \end{aligned} \tag{116}$$

where the Lagrange multiplier $\Phi^{\mu\nu\rho\sigma}$ [58] is a four dimensional tensor density of weight one, symmetric in the index pairs $(\mu\nu)$ and $(\rho\sigma)$ and antisymmetric in each index pair. Thus, Φ has $(6 \cdot 7)/2 = 21$ independent components of which the totally skew component is projected out in (116), leaving us with $36 - 16 = 20$ independent components. Hence the Euler Lagrange equations for Φ precisely delete the amount of unwanted degrees of freedom in B and impose the simplicity constraint. Hence, classically $S'_P[\Omega, B, \Phi]$ and $S_P[\Omega, e]$ are equivalent. Thus, if we write a path integral for S'_P and treat the Lagrange multiplier term S_I in (116) as an interaction Lagrangean (a perturbation) to BF-theory, then we can make use of the powerful techniques that have been developed for the path integral quantization for BF-theory and its perturbation theory.

[10] Another solution is $B^{IJ} = e^I \wedge e^J$ but this possibility is currently not discussed.

EXERCISE 25.

i) Write the Euler Lagrange equations for BF-theory and conclude that the solutions consist of flat connections ω and gauge invariant $B-$ fields. Conclude that ω can be gauged to *zero* by $SL(2,\mathbb{C})$ transformations locally and that then B is closed, that is, locally exact by Poincaré's theorem. Now, verify that the BF-action is not only invariant under local $SL(2,\mathbb{C})$-transformations but also under

$$B^{IJ} =\mapsto B^{IJ} + (D \wedge \theta)^{IJ} = B^{IJ} + d\theta^{IJ} + \omega^I{}_K \wedge \theta^{KJ} + \theta^{IK} \wedge \omega_K{}^J \tag{117}$$

for some $sl(2,\mathbb{C})$ valued one-form θ and that therefore also B can be gauged to zero locally.
Hint: Use the Bianchi identity for Ω.

ii) Perform the Legendre transformation and conclude that there are as many first class constraints as canonical pairs so that again at most a countable number of global degrees of freedom can exist.

One may wonder how it is possible that a theory with less kinematical degrees of freedom has more dynamical (true) degrees of freedom. The answer is that BF-theory has by far more symmetries than the Palatini theory, thus when constraining the number of degrees of freedom we are freezing more symmetries than we deleted degrees of freedom.

Let us now discuss how one formulates the path integral corresponding to the action (116). It is formally given by

$$K_P(\Sigma_{t_1}, \Sigma_{t_2}) = \int [D\omega \; DB \; D\Phi] e^{iS'_P[\omega,B,\Phi]} \,, \tag{118}$$

where $\Sigma_{t_1}, \Sigma_{t_2}$ denote suitable boundary conditions specified in more detail below. Suppose we set $\Phi = 0$, then (118) is a path integral for BF-theory and the integral over B results in the functional δ-distribution $\delta[\Omega]$ imposing the flatness of ω. Now flatness of a connection is equivalent to trivial holonomy along contractible loops by the Ambrose–Singer theorem. If one regularizes the path integral by introducing a triangulation τ of M, then $\delta[F]$ can be written as $\prod_\alpha \delta(\omega(\alpha), 1)$ where the product is over a generating system of independent, contractible loops in τ and $\delta(\omega(\alpha), 1)$ denotes the δ-distribution on $SL(2,\mathbb{C})$ with respect to the Haar measure. Since $SL(2,\mathbb{C})$ is a non-compact group, the δ-distribution is a direct integral over irreducible, unitary representations rather than a direct sum as it would be the case for compact groups (Peter&Weyl theorem). Such representations are infinite dimensional and are labelled by a continuous parameter $\rho \in \mathbb{R}_0^+$ and a discrete parameter $n \in \mathbb{N}_0^+$. Thus, one arrives at a triangulated spin foam model: For a fixed triangulation one integrates (sums) over all possible "spins" ρ (n) that label the generating set of loops (equivalently: the faces that they enclose) of that triangulated four manifold. The analogy with the state sum models for TQFT's is obvious.

Now what one does is a certain jump, whose physical implication is still not understood: Instead of performing perturbation theory in S_I one argues that formally integrating over Φ and thus imposing the simplicity constraint is equivalent to the restriction of the direct integral that enters the δ-distributions to simple representations, that is, representations for which either $n = 0$ or $\rho = 0$. In other words, one says that the triangulated Palatini path integral is the same as the triangulated BF path integral restricted to simple representations. To motivate this argument, one notices that upon canonical quantization of BF theory on a triangulated manifold the B field is the momentum conjugate to ω and if one quantizes on a Hilbert space based on $sl(2,\mathbb{C})$ connections using the Haar measure (similar as we have done for $SU(2)$ for a fixed graph), its corresponding flux operator $\hat{B}_{IJ}(S)$ becomes a linear combination of right invariant vector fields R^{IJ} on $SL(2,\mathbb{C})$. Now the simplicity constraint becomes the condition that the second Casimir operator $R^{IJ}R^{KL}\epsilon_{IJKL}$ vanishes. However, on irreducible representations this operator is diagonal with eigenvalues $n\rho/4$. While this is a strong motivation, it is certainly not sufficient justification for this way of implementing the simplicity constraint in the path integral because it is not clear how this is related to integrating over Φ.

In any case, if one does this then one arrives at (some version of) the Lorentzian Barrett–Crane model [59]. Surprisingly, for a large class of triangulations τ the amplitudes

$$K^\tau_P(\Sigma_{t_1}, \Sigma_{t_2}) := [\int [D\omega\; DB] e^{iS_{BF}[\omega,B]}]_{|\mathrm{simplereps.}} \tag{119}$$

actually converge although one integrates over a non-compact group! This is a non-trivial result [60]. The path integral is then over all possible representations that label the faces of a spin foam and the boundary conditions keep the representations on the boundary graphs, that is, spin networks fixed ($SL(2,\mathbb{C})$ reduces to the $SU(2)$ on the boundary). This also answers the question of what the boundary conditions should be.

There is still an open issue, namely how one should get rid of the regulator (or triangulation) dependence. Since BF-theory is a topological QFT, the amplitudes are automatically triangulation independent, however, this is certainly not the case with GR. One possibility is to sum over triangulations and a concrete proposal of how to weigh the contributions from different triangulations comes from the so-called field theory formulation of the theory [61]. Here one reformulates the BF-theory path integral as the path integral for a scalar field on a group manifold which in this case is a certain power of $SL(2,\mathbb{C})$. The action for that scalar field has a free piece and an interaction piece and performing the perturbation theory (Feynman graphs!) for that field theory is equivalent to the sum over BF-theory amplitudes for all triangulated manifolds *with precisely defined weights*. This idea can be straightforwardly applied also to our context where the restriction to simple representations is realized by imposing corresponding restrictions (projections) on the scalar field.

Summarizing, spin foam models are a serious attempt to arrive at a covariant formulation of QGR but many issues are still unsettled, e.g.:

1. There is no clean equivalence with the Hamiltonian formulation as we have seen. Without that it is unclear how to interpret the spin foam model amplitude and whether it has the correct classical limit. In order to make progress

on the issue of the classical limit, model independent techniques for constructing "causal spin foams" [62] with a built in notion of quantum causality and renormalization methods [63], which should allow in principle the derivation of a low energy effective action, have been developed.

2. The physical correctness of the Barrett–Crane model is unclear. This is emphasized by recent results within the Euclidean formulation [64] which suggest that the classical limit is far off GR since the amplitudes are dominated by spin values close to zero. This was to be expected because in the definition of the Barrett–Crane model there is a certain flexibility concerning the choice of the measure that replaces $[D\omega\ DB]$ at the triangulated level and the result [64] indicates that one must gain more control on that choice.
3. It is not even clear that these models are four-dimensionally covariant: One usually defines that the amplitudes for a fixed triangulation are the same for any four - diffeomorphic triangulation. However, recent results [65] show that this natural definition could result nevertheless in anomalies. This problem is again related to the choice of the measure just mentioned.

Thus, substantially more work is required in order to fill in the present gaps but the results already obtained are very promising indeed.

3.4 Quantum Black Holes

Isolated Horizons. Any theory of quantum gravity must face the question whether it can reproduce the celebrated result due to Bekenstein and Hawking [66] that a black hole in a spacetime (M, g) should account for a quantum statistical entropy given by

$$S_{BH} = \frac{\mathrm{Ar}(H)}{4\ell_{\mathrm{P}}^2}\,, \tag{120}$$

where H denotes the two-dimensional event horizon of the black hole. This result was obtained within the framework of QFT on Curved SpaceTimes (CST) and should therefore be valid in a semiclassical regime in which quantum fluctuations of the gravitational field are negligible (large black holes). The most important question from the point of view of a microscopical theory of quantum gravity is, *what are the microscopical degrees of freedom that give rise to that entropy.* In particular, how can it be within a quantum field theory with an infinite number of degrees of freedom, that this entropy, presumably a measure for our lack of information of what happens behind the horizon, comes out finite.

In [67] the authors performed a bold computation: For *any* surface S and any positive number A_0 they asked the question how many SNW states there are in QGR such that the area operators eigenvalues lie within the interval $[A_0 - \ell_{\mathrm{P}}^2, A_0 + \ell_{\mathrm{P}}^2]$. This answer is certainly infinite because a SNW can intersect S in an uncountably infinite number of different positions without changing the eigenvalues. This divergence can be made less severe by moding out by spatial diffeomorphisms which we can use to map these different SNW onto each other in the vicinity of the surface. However, since there are still an infinite number of non-spatially diffeomorphic states which look the same in the vicinity of the surface but different away from it, the answer is still divergent. Therefore, one has to argue that one must not count information off the surface, maybe invoking the Hamiltonian constraint or using the information that $S = H$ is not an arbitrary surface but actually the horizon of a black hole. Given this assumption, the result of the, actually quite simple counting problem came rather close to (120) with the correct factor of 1/4.

Thus the task left is to justify the assumptions made and to make the entropy counting water-tight by invoking the information that H is a black hole horizon. The outcome of this analysis created a whole industry of its own, known under the name "isolated horizons", which to large part is a beautiful new chapter in classical GR. In what follows we will focus only on a tiny fraction of the framework, mostly concentrating on the ingredients essential for the quantum formulation. For reviews see [68] which also contain a complete list of references on the more classical aspects of this programme, the pivotal papers concerning the quantum applications are [69].

By definition, an event horizon is the external boundary of the part of M that does not lie in the past of null future infinity J^+ in a Penrose diagramme. From an operational point of view, this definition makes little sense because in order to determine whether a candidate is an event horizon, one must know the whole spacetime (M, g) which is never possible by measurements which are necessarily local in spacetime (what looks like an eternal black hole now could capture some dust later and the horizon would change its location). Thus one looks for some local substitute of the notion of an event horizon which captures the idea that the black hole has come to some equilibrium state at least for some amount of time. This is roughly what an isolated horizon Δ is, illustrated in Fig. 14.

More in technical details we have the following.

Definition 13.
A part Δ of the boundary ∂M of a spacetime (M, g) is called an isolated horizon, provided that

1) $\Delta \equiv \mathbb{R} \times S^2$ is a null hypersurface and has zero shear and expansion[11].
2) The field equations and matter energy conditions hold at Δ.
3) g is Lie derived by the null generator l of Δ at Δ.

The canonical formulation of a field theory on a manifold M with boundary Δ must involve boundary conditions at Δ in order that the variation principle be well-defined (the action must be functionally differentiable). Such boundary conditions usually give birth to boundary degrees of freedom [70] which would normally be absent but now come into being because (part of the) gauge transformations are forced to become trivial at Δ. In the present situation what happens

[11] Recall the notions of shear, expansion and twist of a congruence of vector fields in connection with Raychaudhuri's equation.

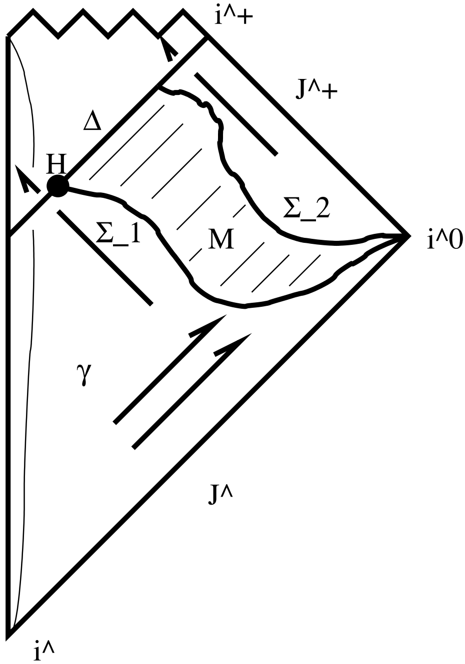


Fig. 14. An isolated horizon Δ boundary of a piece M (shaded) of spacetime also bounded by spacelike hypersurfaces Σ_1, Σ_2. Radiation γ may enter or leave M and propagate into the singularity before or after the isolated horizon has formed but must not cross Δ. An intersection of a spacelike hypersurface Σ with Δ is denoted by H which has the topology of a sphere

is that the boundary term is actually a $U(1)$ *Chern-Simons action*[12]

$$S_{CS} = \frac{A_0}{\pi\beta} \int_\Delta W \wedge dW = \int_{\mathbb{R}} dt \int_H d^2y \epsilon^{IJ} [\dot{W}_I W_J + W_t (dW)_{IJ}] \,, \tag{121}$$

where W is a $U(1)$ connection one form and $H = S^2 = \Sigma \cap \Delta$ is a sphere. The relation between the bulk fields A_a^j, E_j^a and the boundary fields $W_I,\ I = 1,2$ is given by

$$X_H^* A^j = W r^j \text{ and } [X_H^*(*E)_j] r_j = -\frac{A_0}{2\pi\beta} dW \,, \tag{122}$$

where $X_H : \ H \to \Sigma$ is the embedding of the boundary H of Σ into Σ and r^j is an arbitrary but fixed unit vector in $su(2)$ which is to be preserved under $SU(2)$ gauge transformations at Δ and therefore reduces $SU(2)$ to $U(1)$. The number A_0 is the area of H as measured by g which turns out to be a constant

[12] It was observed first in [71] that general relativity in terms of connection variables and in the presence of boundaries leads to Chern–Simons boundary terms.

of the motion as a consequence of the field equations. The existence of r^j is a consequence of definition (13) and $*E$ is the natural metric independent two-form dual to E.

Entropy Counting. One now has to quantize the system. This consists of several steps whose details are complicated and which we will only sketch in what follows.

i) *Kinematical Hilbert Space*
The bulk and boundary degrees of freedom are independent of each other, therefore we choose $\mathcal{H}_{\text{kin}} = \mathcal{H}^{\Sigma}_{\text{kin}} \otimes \mathcal{H}^{H}_{\text{kin}}$ where both spaces are of the form $L_2(\overline{\mathcal{A}}, d\mu_0)$ just that the first factor is for an $SU(2)$ bundle over Σ while the second is for an $U(1)$ bundle over H.

ii) *Quantum Boundary conditions*
Equation (122) implies, in particular, that in quantum theory we must have schematically

$$\left[\widehat{[X^*_H(*E)_j]r_j}\right] \otimes \mathrm{id}_{\mathcal{H}^H_{\text{kin}}} = \mathrm{id}_{\mathcal{H}^\Sigma_{\text{kin}}} \otimes \left[-\frac{A_0}{2\pi\beta}\widehat{dW}\right] . \tag{123}$$

Now we have seen in the bulk theory that we have discussed in great detail throughout this review, that $*E$ is an operator valued distribution which must be smeared by two-surfaces in order to arrive at the well-defined electric fluxes. Since (123) is evaluated at H, this flux operator will non-trivially act only on SNWF's T_s which live in the bulk but intersect H in punctures $p \in H \cap \gamma(s)$. Now the distributional character of the electric fluxes implies that the left hand side of (123) is non-vanishing only at those punctures. Thus the curvature of W is flat everywhere except for the punctures where it is distributional.
Consider now SNWF's T_s of the bulk theory and those of the boundary theory T'_c. Then $\widehat{[X^*_H(*E)_j]r_j]}$ acts on T_s like the z-component of the angular momentum operator and will have distributional eigenvalues proportional to the magnetic quantum numbers m_e of the edges with punctures $p = e \cap H$ and spin j_e where $m_e \in \{-j_e, -j_e+1, .., j_e\}$.

iii) *Implementation of Quantum Dynamics at Δ*
It turns out that $\widehat{X^*_H(*E)_j]r_j}$ and $\widehat{dW}$ are the generators of residual $SU(2)$ gauge transformations close to $X_H(H)$ and of $U(1)$ on H respectively. Now these residual $SU(2)$ transformations are frozen to $U(1)$ transformations by r_j and the most general situation in order for a state to be gauge invariant is that these residual $SU(2)$ transformations of the bulk theory and the $U(1)$ transformations of the boundary theory precisely cancel each other. It turns out that this cancellation condition is precisely given by the quantum boundary condition (123). Thus the states that solve the Gauss constraint are linear combinations of states of the form $T_s \otimes T'_c$ where the boundary data of these states are punctures $p \in \mathcal{P}$ where $p \in \gamma(s) \cap H$, the spins $j_p = j_{e_p}$ of edges $e \in E(\gamma(s))$ with $e_p \cap H = p$ and their magnetic quantum

numbers $m_p = m_{e_p}$. However, due to the specific features of the geometrical quantization of Chern-Simons theories [72] the m_p cannot be specified freely, they have to satisfy the constraint

$$\sum_{p \in \mathcal{P}} 2m_p = 0 \bmod k, \quad k = \frac{A_0}{4\pi \ell_{\mathrm{P}}^2}\,, \tag{124}$$

where k is called the level of a Chern Simons theory which is constrained to be an integer due to Weil's quantization obstruction cocycle criterion of geometric quantization [30] and comes about as follows: The T'_c are actually fixed to be Θ-functions of level k labelled by integers a_p which satisfy the gauge invariance condition

$$2m_p = -a_p \bmod k, \ \sum_p a_p = 0 \bmod k. \tag{125}$$

Next, the spatial diffeomorphism constraint of the bulk theory tells us that the position of the punctures on H are not important, important is only their number.

Finally, one of the boundary conditions at Δ implies that the lapse becomes trivial $N=0$ at H if $\hat{C}(N)$ is to generate an infinitesimal time reparameterization[13]. Thus, luckily we can escape the open issues with the Hamiltonian constraint as far as the quantum dynamics at H is concerned.

We can now come to the issue of entropy counting. First of all we notice that $\mathrm{Ar}(H)$ is a Dirac observable because H is invariant under $\mathrm{Diff}(H)$ and $N=0$ at H. Given n punctures with spins j_l, $l=1,..,N$ the area eigenvalue for H is

$$\lambda(n, \boldsymbol{j}) = 8\pi \ell_{\mathrm{P}}^2 \beta \sum_{n=1}^{n} \sqrt{j_l(j_l+1)}. \tag{126}$$

Now the physical Hilbert space is of the form

$$\mathcal{H}_{\mathrm{phys}} = \oplus_{n,\boldsymbol{j},\boldsymbol{m},\boldsymbol{a}=_k-2\boldsymbol{m}} \ \mathcal{H}^B_{n,\boldsymbol{j},\boldsymbol{m}} \otimes \mathcal{H}^{BH}_{n,\boldsymbol{j},\boldsymbol{m}} \otimes \mathcal{H}^H_{n,\boldsymbol{a}}\,, \tag{127}$$

where $\mathcal{H}^{BH}_{n,\boldsymbol{j},\boldsymbol{m}}$ describes bulk degrees of freedom10.eps at H corresponding to the black hole (finite dimensional), $\mathcal{H}^B_{n,\boldsymbol{j},\boldsymbol{m}}$ describes bulk degrees of freedom away from H and finally $\mathcal{H}^H_{n,\boldsymbol{a}}$ describes Chern–Simons degrees of freedom which are completely fixed in terms of $\boldsymbol{m}$ due to reasons of gauge invariance (125). The situation is illustrated in Fig. 15. Let $\delta > 0$ and let $S_{A_0,\delta}$ be the set of eigenstates

[13] This does not mean that the lapse of a classical isolated horizon solution must vanish at S, rather there is a subtle difference between gauge motions and symmetries for field theories with boundaries [70] where in this case symmetries map solutions to gauge inequivalent or equivalent ones respectively, if $N_{|H} \neq 0$ or $N_{|H} = 0$ respectively.

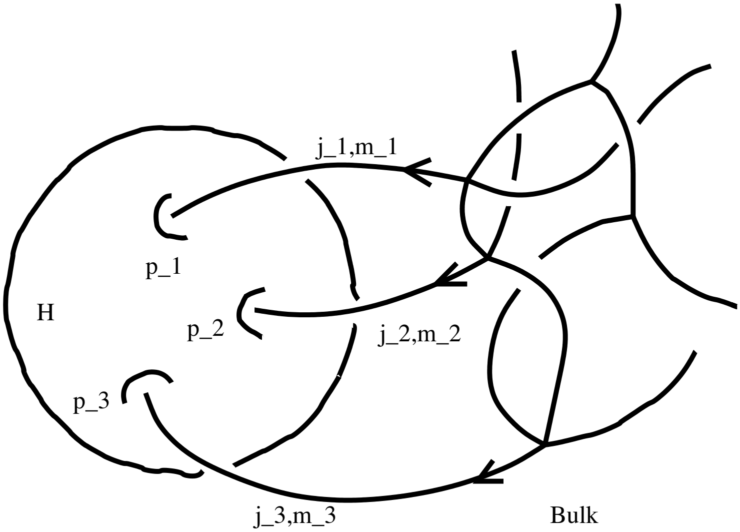


Fig. 15. Punctures, spins, magnetic quantum numbers and entropy counting. Only the relevant boundary data are shown, the bulk information is traced over

$\psi_{n,\boldsymbol{j},\boldsymbol{m}} \in \mathcal{H}^{BH}_{n,\boldsymbol{j},\boldsymbol{m}}$ of the area operator such that the eigenvalue lies in the interval $[A_0 - \delta, A_0 + \delta]$ and $N_{A_0,\delta}$ their number. Define the density matrix

$$\hat{\rho}_{BH} = \mathrm{id}_B \otimes \left[\frac{1}{N_{A_0,\delta}} \sum_{\psi \in S_{A_0,\delta}} |\psi >< \psi| \right] \otimes \mathrm{id}_H . \tag{128}$$

The quantum statistical entropy from this microcanonical ensemble is given by

$$S_{BH} = -\mathrm{Tr}(\hat{\rho}_{BH} \ln(\hat{\rho}_{BH})) = \ln(N_{A_0,\delta}) . \tag{129}$$

Thus we just need to count states and the answer *will be finite because the area operator has an area gap.*

Exercise 26.
Estimate $N_{A_0,\delta}$ from above and below taking into account the constraint (124) and that k is an integer (purely combinatorial problem!).

The result of the counting problem is that S_{BH} is indeed given by (120) to leading order in A_0 (there are logarithmic corrections) for $\delta \approx \ell_{\mathrm{P}}^2$ provided that

$$\beta = \frac{\ln(2)}{\pi\sqrt{3}} . \tag{130}$$

Here the numbers $\ln(2), \sqrt{3}$ comes from the fact that the configurations with lowest spin $j_l = 1/2$ make the dominant contribution to the entropy with eigenvalue $\propto \beta n \sqrt{3} \approx A_0$ and number of states given by $N_{A_0,\delta} \approx 2^n$ that is, *two Boolean* *degrees of freedom per puncture* [73]. This provides an explicit explanation for the origin of the entropy. Now fixing β at the value (130) would make little sense would it be different for different types of black hole (that is, in presence of different matter, charges, rotation, other hair,..). However, this is not the case!

In summary, the analysis sketched above provides a self-contained derivation of S_{BH} within QGR. The result is *highly non-trivial* because it was not to be expected from the outset that Loop Quantum Gravity, classical GR and Chern Simons theory would interact in such a harmonic way as to provide the expected result: Chern–Simons theory is *very different* from QGR and still they have an interface at H. The result applies to astrophysically interesting black holes of the Schwarzschild type and does not require supersymmetry. Nevertheless, the calculation still has a semiclassical input because the presence of the isolated horizon is fed in at the classical level already. It would be more satisfactory to have a *quantum definition of an (isolated) horizon* but this is a hard task and left for future research. Another unsolved problem then is the calculation of the Hawking effect from first principles.

3.5 Semiclassical Analysis

The Complexifier Machinery for Generating Coherent States. Let us first specify what we mean by semiclassical states.

Definition 14.
Let be given a phase space $\mathcal{M},\ \{.,.\}$ with preferred Poisson subalgebra $\mathcal{O}$ of $C^\infty(\mathcal{M})$ and a Hilbert space $\mathcal{H},\ [.,.]$ together with an operator subalgebra $\hat{\mathcal{O}}$ of $\mathcal{L}(\mathcal{H})$. The triple $\mathcal{M},\ \{.,.\},\mathcal{O}$ is said to be a classical limit of the triple $\mathcal{H},\ [.,.],\hat{\mathcal{O}}$ provided that there exists an (over)complete set of states $\{\psi_m\}_{m\in\mathcal{M}}$ such that for all $O,O'\in\mathcal{O}$ the *infinitesimal Ehrenfest property*

$$\left|\frac{<\hat{O}>_m}{O(m)}-1\right|\ll 1\,,\qquad \left|\frac{<[\hat{O},\hat{O}']>_m}{i\hbar\{O,O'\}(m)}-1\right|\ll 1 \tag{131}$$

and the *small fluctuation property*

$$\left|\frac{<\hat{O}^2>_m}{<\hat{O}>_m^2}-1\right|\ll 1 \tag{132}$$

holds at generic[14] points in $\mathcal{M}$. Here $<\,.\,>_m:=<\psi_m,.\psi_m>/||\psi_m||^2$ denotes the expectation value functional.

For systems with constraints, strictly speaking, semiclassical states should be physical states, that is, those that solve the constraints because we are not interested in approximating gauge degrees of freedom but only physical observables. Only then are the predictions ($\hbar$ corrections to the classical limit) of the

[14] The set of points where (131), (132) are violated should have small Liouville measure.

theory reliable. In the present situation with QGR, however, we are more interested in constructing *kinematical semiclassical states* for the following reason: As we have shown, the status of the physical correctness of the Hamiltonian constraint operator $\hat{C}$ is unsettled. We would therefore like to test whether it has the correct classical limit. This test is obviously meaningless on states which the Hamiltonian constraint annihilates anyway. For the same reason it also does not make sense to construct semiclassical states which are at least spatially diffeomorphism invariant because the Hamiltonian constraint does not leave this space invariant.

The key question then is how to construct semiclassical states. Fortunately, for phase spaces which have a cotangent bundle structure as is the case with QGR, a rather general construction guideline is available [74], the so-called *Complexifier Method*, which we will now sketch:

Let $(\mathcal{M}, \{.,.\})$ be a phase space with (strong) symplectic structure $\{.,.\}$ (notice that $\mathcal{M}$ is allowed to be infinite dimensional). We will assume that $\mathcal{M} = T^*\mathcal{C}$ is a cotangent bundle. Let us then choose a real polarization of $\mathcal{M}$, that is, a real Lagrangean submanifold $\mathcal{C}$ which will play the role of our configuration space. Then a loose definition of a complexifier is as follows:

Definition 15.
A complexifier is a positive definite function [15] C on $\mathcal{M}$ with the dimension of an action, which is smooth a.e. (with respect to the Liouville measure induced from $\{.,.\}$) and whose Hamiltonian vector field is everywhere non-vanishing on $\mathcal{C}$. Moreover, for each point $q \in \mathcal{C}$ the function $p \mapsto C_q(p) = C(q,p)$ grows stronger than linearly with $||p||_q$ where p is a local momentum coordinate and $||.||_q$ is a suitable norm on $T^*_q(\mathcal{C})$.

In the course of our discussion we will motivate all of these requirements.

The reason for the name *complexifier* is that C enables us to generate a *complex polarization* of $\mathcal{M}$ from $\mathcal{C}$ as follows: If we denote by q local coordinates of $\mathcal{C}$ (we do not display any discrete or continuous labels but we assume that local fields have been properly smeared with test functions) then

$$z(m) := \sum_{n=0}^{\infty} \frac{i^n}{n!} \{q, C\}_{(n)}(m) \tag{133}$$

define local complex coordinates of $\mathcal{M}$ provided we can invert $z, \bar{z}$ for $m := (q,p)$ where p are the fibre (momentum) coordinates of $\mathcal{M}$. This is granted at least locally by definition 15. Here the multiple Poisson bracket is inductively defined by $\{q,C\}_{(0)} = q$, $\{q,C\}_{(n+1)} = \{\{q,C\}_{(n)}, C\}$ and makes sense due to the required smoothness. What is interesting about (133) is that it implies the following bracket structure

$$\{z, z\} = \{\bar{z}, \bar{z}\} = 0 \tag{134}$$

[15] For the rest of this section C will denote a complexifier function and not the Hamiltonian constraint.

while $\{z, \bar{z}\}$ is necessarily non-vanishing. The reason for this is that (133) may be written in the more compact form

$$z = e^{-i\mathcal{L}_{\chi_C}} q = ([\varphi^t_{\chi_C}]^* q)_{t=-i} \,, \tag{135}$$

where χ_C denotes the Hamiltonian vector field of C, $\mathcal{L}$ denotes the Lie derivative and $\varphi^t_{\chi_C}$ is the one-parameter family of canonical transformations generated by χ_C. Formula (135) displays the transformation (133) as the analytic extension to imaginary values of the one parameter family of diffeomorphisms generated by χ_C and since the flow generated by Hamiltonian vector fields leaves Poisson brackets invariant, (134) follows from the definition of a Lagrangean submanifold. The fact that we have continued to the negative imaginary axis rather than the positive one is important in what follows and has to do with the required positivity of C.

The importance of this observation is that either of $z, \bar{z}$ are coordinates of a Lagrangean submanifold of the complexification $\mathcal{M}^{\mathbb{C}}$, i.e. a complex polarization and thus may serve to define a Bargmann-Segal representation of the quantum theory (wave functions are holomorphic functions of z). The diffeomorphism $\mathcal{M} \to \mathcal{C}^{\mathbb{C}}$; $m \mapsto z(m)$ shows that we may think of $\mathcal{M}$ either as a symplectic manifold or as a complex manifold (complexification of the configuration space). Indeed, the polarization is usually a positive Kähler polarization with respect to the natural $\{.,.\}$-compatible complex structure on a cotangent bundle defined by local Darboux coordinates, if we choose the complexifier to be a function of p only. These facts make the associated Segal-Bargmann representation especially attractive.

We now apply the rules of canonical quantization: a suitable Poisson algebra $\mathcal{O}$ of functions O on $\mathcal{M}$ is promoted to an algebra $\hat{\mathcal{O}}$ of operators $\hat{O}$ on a Hilbert space $\mathcal{H}$ subject to the condition that Poisson brackets turn into commutators divided by $i\hbar$ and that reality conditions are reflected as adjointness relations, that is,

$$[\hat{O}, \hat{O}'] = i\hbar \widehat{\{O, O'\}} + o(\hbar), \quad \hat{O}^\dagger = \hat{\bar{O}} + o(\hbar) \,, \tag{136}$$

where quantum corrections are allowed (and in principle unavoidable except if we restrict $\mathcal{O}$, say to functions linear in momenta). We will assume that the Hilbert space can be represented as a space of square integrable functions on (a distributional extension $\overline{\mathcal{C}}$ of) $\mathcal{C}$ with respect to a positive, faithful probability measure μ, that is, $\mathcal{H} = L_2(\overline{\mathcal{C}}, d\mu)$ as it is motivated by the real polarization.

The fact that C is positive motivates to quantize it in such a way that it becomes a self-adjoint, positive definite operator. We will assume this to be the case in what follows. Applying then the quantization rules to the functions z in (133) we arrive at

$$\hat{z} = \sum_{n=0}^{\infty} \frac{i^n}{n!} \frac{[\hat{q}, \hat{C}]_{(n)}}{(i\hbar)^n} = e^{-\hat{C}/\hbar} \hat{q} e^{\hat{C}/\hbar}. \tag{137}$$

The appearance of $1/\hbar$ in (137) justifies the requirement for $C/\hbar$ to be dimensionless in definition 15. We will call $\hat{z}$ *annihilation operator* for reasons that will become obvious in a moment.

Let now $q \mapsto \delta_{q'}(q)$ be the δ-distribution with respect to μ with support at $q = q'$. (More in mathematical terms, consider the complex probability measure, denoted as $\delta_{q'} d\mu$, which is defined by $\int \delta_{q'} d\mu f = f(q')$ for measurable f). Notice that since C is non-negative and necessarily depends non-trivially on momenta (which will turn into (functional) derivative operators in the quantum theory), the operator $e^{-\hat{C}/\hbar}$ is a *smoothening operator*. Therefore, although $\delta_{q'}$ is certainly not square integrable, the complex measure (which is probability if $\hat{C} \cdot 1 = 0$)

$$\psi_{q'} := e^{-\hat{C}/\hbar} \delta_{q'} \tag{138}$$

has a chance to be an element of $\mathcal{H}$. Whether or not it does depends on the details of $\mathcal{M}, \{.,.\}, C$. For instance, if C as a function of p at fixed q has flat directions, then the smoothening effect of $e^{-\hat{C}/\hbar}$ may be insufficient, so in order to avoid this we required that C is positive definite and not merely non-negative. If C would be indefinite, then (138) has no chance to make sense as an L_2 function.

We will see in a moment that (138) qualifies as a candidate *coherent state* if we are able to analytically extend (138) to complex values z of q' where the label z in ψ_z will play the role of the point in $\mathcal{M}$ at which the coherent state is peaked. In order that this is possible (and in order that the extended function is still square integrable), (138) should be entire analytic. Now $\delta_{q'}(q)$ roughly has an integral kernel of the form $e^{i(k,(q-q'))}$ (with some pairing $< .,. >$ between tangential and cotangent vectors) which is analytic in q' but the integral over k, after applying $e^{-\hat{C}/\hbar}$, will produce an entire analytic function only if there is a damping factor which decreases faster than exponentially. This provides the intuitive explanation for the growth requirement in definition 15. Notice that the ψ_z are not necessarily normalized.

Let us then assume that

$$q \mapsto \psi_m(q) := [\psi_{q'}(q)]_{q' \to z(m)} = [e^{-\hat{C}/\hbar} \delta_{q'}(q)]_{q' \to z(m)} \tag{139}$$

is an entire L_2 function. Then ψ_m is automatically an *eigenfunction of the annihilation operator $\hat{z}$ with eigenvalue z* since

$$\hat{z}\psi_m = [e^{-\hat{C}/\hbar} \hat{q} \delta_{q'}]_{q' \to z(m)} = [q' e^{-\hat{C}/\hbar} \delta_{q'}]_{q' \to z(m)} = z(m) \psi_m \,, \tag{140}$$

where in the second step we used that the delta distribution is a generalized eigenfunction of the operator $\hat{q}$. But to be an eigenfunction of an annihilation operator *is one of the accepted definitions of coherent states!*

Next, let us verify that ψ_m indeed has a chance to be peaked at m. To see this, let us consider the self-adjoint (modulo domain questions) combinations

$$\hat{x} := \frac{\hat{z} + \hat{z}^\dagger}{2}, \quad \hat{y} := \frac{\hat{z} - \hat{z}^\dagger}{2i} \tag{141}$$

whose classical analogs provide real coordinates for $\mathcal{M}$. Then we have automatically from (140)

$$< \hat{x} >_m := \frac{< \psi_m, \hat{x}\psi_m >}{||\psi_m||^2} = \frac{z(m) + \bar{z}(m)}{2} =: x(m) \tag{142}$$

and similar for y. Equation (142) tells us that the operator $\hat{z}$ should really correspond to the function $m \mapsto z(m),\ m \in \mathcal{M}$.

Now we compute by similar methods that

$$< [\delta\hat{x}]^2 >_m := \frac{< \psi_m, [\hat{x} - < \hat{x} >_m]^2 \psi_m >}{||\psi_m||^2} = < [\delta\hat{y}]^2 >_m = \frac{1}{2} | < [\hat{x}, \hat{y}] >_m | , \tag{143}$$

so that the ψ_m are automatically *minimal uncertainty states for* $\hat{x}, \hat{y}$, moreover the fluctuations are unquenched (equal each other). This is the second motivation for calling the ψ_m coherent states. Certainly one should not only check that the fluctuations are minimal but also that they are small as compared to the expectation value, at least at generic points of the phase space, in order that the quantum errors are small.

The *infinitesimal Ehrenfest* property

$$\frac{< [\hat{x}, \hat{y}] >_z}{i\hbar} = \{x, y\}(m) + O(\hbar) \tag{144}$$

follows if we have properly implemented the canonical commutation relations and adjointness relations. The size of the correction, however, does not follow from these general considerations but the minimal uncertainty property makes small corrections plausible. Condition (144) supplies information about how well the symplectic structure is reproduced in the quantum theory.

For the same reason one expects that the peakedness property

$$|\frac{< \psi_m, \psi_{m'} > |^2}{||\psi_m||^2 \ ||\psi_{m'}||^2} \approx \chi_{K_m}(m') \tag{145}$$

holds, where K_m is a phase cell with center m and Liouville volume $\approx \sqrt{< [\delta\hat{x}]^2 >_m < [\delta\hat{y}]^2 >_m}$ and χ denotes the characteristic function of a set.

Finally one wants coherent states to be overcomplete in order that every state in $\mathcal{H}$ can be expanded in terms of them. This has to be checked on a case by case analysis but the fact that our complexifier coherent states are for real z nothing else than regularized δ distributions which in turn provide a (generalized) basis makes this property plausible to hold.

Exercise 27.
Consider the phase space: $\mathcal{M} = T^*\mathbb{R} = \mathbb{R}^2$ with standard Poisson brackets $\{q, q\} = \{p, p\} = 0,\ \{p, q\} = 1$ and configuration space $\mathcal{C} = \mathbb{R}$. Consider the complexifier $C = p^2/(2\sigma)$ where σ is a dimensionful constant such that $C/\hbar$ is dimensionless. Check that it meets all the requirements of definition 15 and perform the coherent state construction displayed above.

Hint: Up to a phase, the resulting, normalized coherent states are the usual ones for the harmonic oscillator with Hamiltonian $H = (p^2/m + m\omega^2 q^2)/2$ with $\sigma = m\omega$. Verify that the states ψ_m are Gaussian peaked in the configuration representation with width $\sqrt{\hbar/\sigma}$ around $q = q_0$ and in the momentum representation around $p = p_0$ with width $\sqrt{\hbar\sigma}$ where $m = (p_0, q_0)$.

As it has become clear from the discussion, the complexifier method gives a rough guideline, but no algorithm, in order to arrive at a satisfactory family of coherent states, there are things to be checked on a case by case basis. On the other hand, what is nice is that given *only one input*, namely the complexifier C, it is possible to arrive at a definite and constructive framework for a semiclassical analysis. It is important to know what the classical limit of $\hat{C}$ is, otherwise, if we have just an abstract operator $\hat{C}$ then the map $m \mapsto z(m)$ is unknown and an interpretation of the states in terms of $\mathcal{M}$ is lost.

Application to QGR. Let us now apply these ideas to QGR. Usually the choice of C is strongly motivated by a Hamiltonian, but in QGR we have none. Therefore, at the moment the best we can do is to play with various proposals for $\hat{C}$ and to explore the properties of the resulting states. For the simplest choice of $\hat{C}$ [75] those properties have been worked out more or less completely and we will briefly describe them below.

The operator $\hat{C}$ is defined by its action on cylindrical functions $f = p_\gamma^* f_\gamma$ by

$$\frac{\hat{C}}{\hbar} f = -p_\gamma^* [\frac{1}{2} [\sum_{e \in E(\gamma)} l(e) [R_e^j/2]^2]\, f_\gamma] =: p_\gamma^* [\hat{C}_\gamma f_\gamma]\,, \tag{146}$$

where the positive numbers $l(e)$ satisfy $l(e \circ e') = l(e) + l(e')$ and $l(e^{-1}) = l(e)$ and R_e^j are the usual right invariant vector fields.

Exercise 28.
Recall the definition of the maps $p_{\gamma'\gamma}$ for $\gamma \prec \gamma'$ from Sect. 2.1 and check that the $\hat{C}_\gamma$ are consistently defined, that is, $\hat{C}_{\gamma'} \circ p_{\gamma'\gamma}^* = p_{\gamma'\gamma}^* \circ \hat{C}_\gamma$.

This choice is in analogy to the harmonic oscillator where the quantum complexifier is essentially the Laplacian $-(d/dx)^2$. The classical limit of (146) depends in detail on the function l which is analogous to the parameter $\hbar/\sigma$ for the case of the harmonic oscillator. For instance [74], one can choose a) three families of foliations $s \mapsto H_s^I,\ I = 1, 2, 3$ of σ by two dimensional surfaces H_s^I such that there is a bijection $(s^1, s^2, s^3) \mapsto x(\boldsymbol{s}) := [\cap_I H^I_{s^I}] \in \sigma$ and b) a partition P_s^I of the H_s^I into small surfaces S and define

$$C = \frac{1}{2a^2\kappa} \int_{\mathbb{R}} ds \sum_{I=1}^{3} \sum_{S \in P_s^I} [\mathrm{Ar}(S)]^2 \tag{147}$$

where Ar(S) is again the area functional and a is a dimensionful constant of dimension cm^1. The function l for this example is then roughly[16] $l(e) = \frac{(\beta\ell_{\mathrm{P}})^2}{a^2}\int ds \sum_I \sum_{S\in P_s^I} \chi_S(e)$ where $\chi_S(e) = 1$ if $S \cap e \neq \emptyset$ and vanishes otherwise.

The δ-distribution with respect to the measure μ_0 can be written as the sum over all SNW's (exercise!)

$$\delta_{A'}(A) = \sum_s T_s(A')\overline{T_s(A)} \tag{148}$$

with resulting coherent states

$$\psi_{A^{\mathbb{C}}}(A) = \sum_s e^{-\frac{1}{2}\sum_{e\in E(\gamma(s))} l(e) j_e(j_e+1)} T_s(A^{\mathbb{C}})\overline{T_s(A)}\,, \tag{149}$$

where the $SL(2,\mathbb{C})$ *connection* $A^{\mathbb{C}}$ is defined by

$$A^{\mathbb{C}}[A,E] = \sum_{n=0}^{\infty} \frac{i^n}{n!}(\{A,C\}_{(n)})[A,E]\,. \tag{150}$$

Thus we see that in this case the symplectic manifold given as the cotangent bundle $\mathcal{M} = T^*\mathcal{A}$ over the space of $SU(2)$ connections is also naturally given as the complex manifold $\mathcal{A}^{\mathbb{C}}$ of $SL(2,\mathbb{C})$ connections. From the general discussion above it now follows that the classical interpretation of the *annihilation operators*

$$\hat{A}^{\mathbb{C}}(e) := e^{-\hat{C}/\hbar}\hat{A}(e)e^{\hat{C}/\hbar} \tag{151}$$

is simply the holonomy of the complex connection $A^{\mathbb{C}}(e)$.

In order to study the semiclassical properties of these states we consider their *cut-offs* $\psi_{A^{\mathbb{C}},\gamma}$ for each graph γ defined on cylindrical functions $f = p_\gamma^* f_\gamma$ by

$$< \overline{\psi_{A^{\mathbb{C}}}}, f >_{\mathrm{kin}} =: < \overline{\psi_{A^{\mathbb{C}},\gamma}}, f >_{\mathrm{kin}}\ . \tag{152}$$

Now, (if we work at the non-gauge invariant level,) one can check that

$$\psi_{A^{\mathbb{C}},\gamma}(A) = \prod_{e\in E(\gamma(s)} \psi^{l(e)}_{A^{\mathbb{C}}(e)}(A(e))\,, \tag{153}$$

where for any $g \in SL(2,\mathbb{C}),\ h \in SU(2)$ we have defined

$$\psi_g^t(h) := \sum_j (2j+1)e^{-tj(j+1)/2}\chi_j(gh^{-1})\,. \tag{154}$$

Exercise 29.
Verify, using the Peter&Weyl theorem, that for $g \in SU(2)$ we have $\psi_g^0(h) = \delta_g(h)$, the

[16] This formula gets exact in the limit of infinitely fine partition, at finite coarseness, it is an approximation to the exact, more complicated formula.

δ-distribution with respect to $L_2(SU(2), d\mu_H)$. Conclude that (154) is just the analytic extension of the heat kernel $e^{-t\Delta/2}$ where Δ is the Laplacian on $SU(2)$. Thus the states (154) are in complete analogy with those for the harmonic oscillator, just that $\mathbb{R}$ was replaced by $SU(2)$ and the complexification $\mathbb{C}$ of $\mathbb{R}$ by the complexification $SL(2,\mathbb{C})$ of $SU(2)$. In this form, coherent states on compact gauge groups were originally proposed by Hall [76].

The analysis of the semiclassical properties of the states $\psi_{A^{\mathbb{C}},\gamma}$ on $\mathcal{H}_{\rm kin}$ can therefore be reduced to that of the states ψ^t_g on $L_2(SU(2), d\mu_H)$. We state here without proof that the following properties could be proved [75]: I) Overcompleteness, II) expectation value property, III) Ehrenfest property, IV) peakedness in phase space, V) annihilation operator eigenstate property, VI) minimal uncertainty property and VII) small fluctuation property. Thus, these states have many of the desired properties that one requires from coherent states.

In the following graphic we display as an example the peakedness properties of the analog of (154) for the simpler case of the gauge group $U(1)$, the case of $SU(2)$ is similar but requires more plots because of the higher dimensionality of $SU(2)$. Thus $g_0 = e^p h_0 \in U(1)^{\mathbb{C}} = \mathbb{C} - \{0\},\ p \in \mathbb{R},\ h_0 \in U(1)$ and $u \in U(1)$ where we parameterize $u = e^{i\phi},\ \phi \in [-\pi, \pi)$. Similarly, $g = e^{p_1} u,\ p_1 \in \mathbb{R}$. We consider in Fig. 16 the peakedness in the configuration representation given by the probability amplitude

$$u = e^{i\phi} \mapsto j^t_{g_0}(u) = |\psi^t_{g_0}(u)|^2 / ||\psi^t_{g_0}||^2 \tag{155}$$

at $h_0 = 1, p \in [-5, 5]$. In Fig. 17 the phase space peakedness expressed by the overlap function

$$g = e^{p_1} u \mapsto i^t(g, g_0) = \frac{| < \psi^t_g, \psi^t_{g_0} > |^2}{||\psi^t_g||^2 \; ||\psi^t_{g_0}||^2} \tag{156}$$

is shown at fixed $p = 0, h_0 = 1$ for $p \in [-5, 5],\ u \in U(1)$. We have made use of the fact (exercise!) that $\psi_{g_0}(u)$ and $< \psi_g, \psi_{g_0} >$ respectively depend only on the

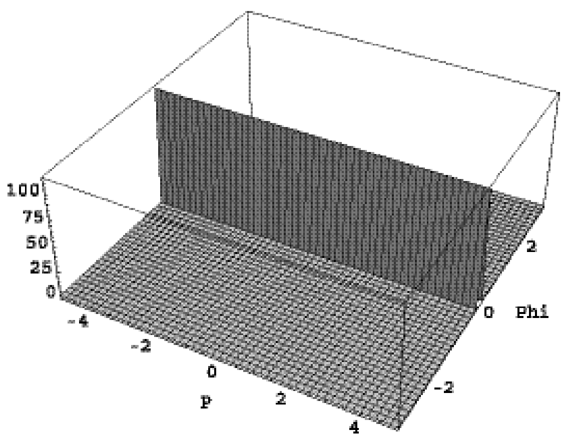


Fig. 16. Probability amplitude $u \mapsto j^t_{g_0}(u)$ at $p \in [-5, 5], h_0 = 1$

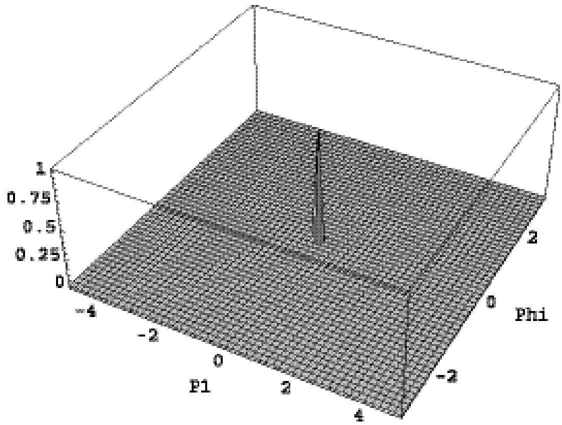


Fig. 17. Overlap function $g \mapsto i^t_{g_0}(g)$ at $p = 0, h_0 = 1$ for $p_1 \in [-5, 5], u \in U(1)$

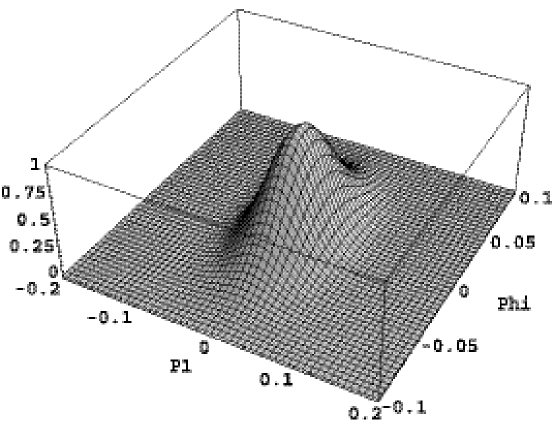


Fig. 18. Resolution of a neighbourhood of the peak of the function $g \mapsto i^t_{g_0}(g)$ at $p = 0, h_0 = 1$

combinations $g_0 u^{-1} = e^{p_0} h u^{-1}$ and $\bar{g} g_0 = e^{p+p_1} u^{-1} h_0$ respectively. Therefore, peakedness at $u = h_0$ or $g = g_0 = e^p h_0$ respectively for any h_0 is equivalent to peakedness at $u = 1$ or at $g = e^{p_0}$ respectively for $h_0 = 1$. Both plots are for the value $t = 0.001$ and one clearly sees the peak width of $\sqrt{t} \approx 0.03$ when resolving those plots around the peak as in Fig. 18, which has a close to Gaussian shape just like the harmonic oscillator coherent states have. As a first modest application, these states have been used in order to analyze how one would obtain, at least in principle, the QFT's on CST's (Curved SpaceTime) limit from full QGR in [49]. In particular, it was possible to perform a detailed calculation concerning the existence of Poincaré invariance violating dispersion relations of photon propagation within QGR which were discussed earlier at a more phenomenological level in the pioneering papers [77]: The idea is that the metric field is a collection of quantum operators which are not mutually commuting. Therefore it should be impossible to construct a state which is peaked on, say the Minkowski metric, and which is a simultaneous eigenstate of all the metric operator components, in other words, there should be no such thing as a Poincaré-invariant state in full QGR[17], already because such an object should be highly background dependent. The best one can construct is a coherent state peaked on the Minkowski metric. The small fluctuations that are encoded in that state influence the propagation of matter and these tiny disturbances could accumulate to measurable sizes in so-called γ-ray burst experiments [78] where one measures the time delay of photons of higher energy as compared to those of lower energy as they travel over cosmological distances as a result of the energy dependence of the speed of light. If such an effect exists then it is a non-perturbative one because perturbatively defined QFT's on Minkowski space are by construction Poincaré invariant (recall e.g. the Wightman axioms from Sect. 1.1).

These are certainly only first moderate steps. The development of the semiclassical analysis for QGR is still in its very beginning and there are many interesting and new mathematical and physical issues that have to be settled before one can seriously attack the proof that, for instance, the Hamiltonian constraint of Sect. 3.1 has the correct classical limit or that full QGR reduces to classical GR plus the standard model in the low energy regime.

3.6 Gravitons

The Isomorphism. The reader with a strong background in ordinary QFT and/or string theory will have wondered throughout these lectures where in QGR the graviton, which plays such a prominent role in the perturbative, background dependent approaches to quantum gravity, resides. In fact, if one understands the graviton, as usually, as an excitation of the quantum metric around Minkowski space, then there is a clear connection with the semiclassical analysis of the previous section: One should construct a suitable coherent state which is peaked on the gauge invariant phase space point characterizing Minkowski space and identify suitable excitations thereof as gravitons. It is clear that at the moment such graviton states from full QGR cannot be constructed, because we would need first to solve the Hamiltonian constraint.

However, one can arrive at an approximate notion of gravitons through the quantization of *linearized gravity*: Linearized gravity is nothing else than the expansion of the full GR action around the gauge variant initial data $(E^0)^a_j = \delta^a_j,\ (A^0)^j_a = 0$ to second order in $E - E^0, A$ which results in a free, classical field theory with constraints. In fact, the usual notion of gravitons is *precisely*

[17] This seems to contradict the fact that we are even interested in four dimensionally diffeomorphism invariant states and the fact that the Poincaré group should be a tiny subgroup thereof. However, this is not the case because we require the states only to be invariant under diffeomorphisms which are pure gauge and those have to die off at spatial infinity. Poincaré transformations are therefore not gauge transformation but symmetries and what we are saying is that there are no Poincaré symmetric, diffeomorphism gauge invariant states.

the ordinary Fock space quantization of that classical, free field theory [79]. In order to see whether QGR can possibly accommodate these graviton states, Varadarajan in a beautiful series of papers [80] has carried out a polymer like quantization of that free field theory on a Hilbert space $\mathcal{H}_{\rm kin}$ which is in complete analogy to that for full QGR, the only difference being that the gauge group $SU(2)$ is replaced by the gauge group $U(1)^3$. While there are certainly large differences between the highly interacting QGR theory and linearized gravity, one should at least be able to gain some insight into the the answer to the question, how a Hilbert space in which the excitations are one dimensional can possibly describe the Fock space excitations (which are three dimensional).

The problem of describing gravitons within linearized gravity by polymer like excitations is mathematically equivalent to the simpler problem of describing the photons of the ordinary Fock Hilbert space $\mathcal{H}_F$ of Maxwell theory by polymer like excitations within a Hilbert space $\mathcal{H}_P = L_2(\overline{\mathcal{A}}, d\mu)$ where $\overline{\mathcal{A}}$ is again a space of generalized $U(1)$ connections with some measure μ thereon. Thus, we describe the latter problem in some detail since it requires less space and has the same educational value.

The crucial observation is the following isomorphism $\mathcal{I}$ between two different Poisson subalgebras of the Poisson algebra on the phase space $\mathcal{M}$ of Maxwell theory coordinatized by a canonical pair (E, A) defined by a $U(1)$ connection A and a conjugate electric field E: Consider a one-parameter family of test functions of rapid decrease which are regularizations of the δ-distribution, for instance

$$f_r(x,y) = \frac{e^{-\frac{||x-y||^2}{2r^2}}}{(\sqrt{2\pi}r)^3}\,, \tag{157}$$

where we have made use of the Euclidean spatial background metric. Given a path $p \in \mathcal{P}$ we denote its distributional *form factor* by

$$X^a_p(x) := \int_0^1 dt\;\; \dot{p}^a(t)\delta(x, p(t))\,. \tag{158}$$

The smeared form factor is defined by

$$X^a_{p,r}(x) := \int d^3y f_r(x,y) X^a_p(y) = \int_0^1 dt\;\; \dot{p}^a(t) f_r(x, p(t))\,, \tag{159}$$

which is evidently a test function of rapid decrease. Notice that a $U(1)$ holonomy maybe written as

$$A(p) := e^{i\int d^3x X^a_p(x) A_a(x)} \tag{160}$$

and we can define a smeared holonomy by

$$A_r(p) := e^{i\int d^3x X^a_{p,r}(x) A_a(x)}\,. \tag{161}$$

Likewise we may define smeared electric fields as

$$E^a_r(x) := \int d^3y f_r(x,y) E^a(y)\,. \tag{162}$$

If we denote by q the electric charge (notice that in our notation $\alpha = \hbar q^2$ is the fine structure constant), then we obtain the following Poisson subalgebras: On the one hand we have smeared holonomies but unsmeared electric fields with

$$\{A_r(p), A_r(p')\} = \{E^a(x), E^b(y)\} = 0, \;\; \{E^a(x), A_r(p)\} = iq^2 X^a_{p,r}(x) A_r(p) \tag{163}$$

and on the other hand we have unsmeared holonomies but smeared electric fields with

$$\{A(p), A(p')\} = \{E^a_r(x), E^b_r(y\} = 0, \;\; \{E^a_r(x), A(p)\} = iq^2 X^a_{p,r}(x) h_p\,. \tag{164}$$

Thus the two Poisson algebras are isomorphic and also the * relations are isomorphic, both $E^a(x), E^a_r(x)$ are real valued while both $A(p), A_r(p)$ are $U(1)$ valued. Thus, as abstract *- Poisson algebras these two algebras are indistinguishable and we may ask if we can find different representations of it. Even better, notice that $A_r(p)A_r(p') = A_r(p \circ p'),\;\; A_r(p)^{-1} = A_r(p^{-1})$ so the smeared holonomy algebra is also isomorphic to the unsmeared one. Hence there is an algebra *-isomorphism $\mathcal{I}$ defined on the generators by $\mathcal{I}_r(h_p) = h_{p,r},\;\; \mathcal{I}_r(E_r) = E$. One must also show that the $A_r(p)$ are still algebraically independent as are the $A(p)$ [80].

Induced Fock Representation with Polymer-Excitations. Now we know that the unsmeared holonomy algebra is well represented on the Hilbert space $\mathcal{H}_{\rm kin} = L_2(\overline{\mathcal{A}}, d\mu_0)$ while the smeared holonomy algebra is well represented on the Fock Hilbert space $\mathcal{H}_F = L_2(\mathcal{S}', d\mu_F)$ where $\mathcal{S}'$ denotes the space of divergence free, tempered distributions and μ_F is the Maxwell-Fock measure. These measures are completely characterized by their generating functional

$$\omega_F(\hat{A}_r(p)) := \mu_F(A_r(p)) = e^{-\frac{1}{4\alpha}\int d^3x X^a_{p,r}(x)\sqrt{-\Delta}^{-1} X^b_{p,r}\delta_{ab}} \tag{165}$$

since finite linear combinations of the $h_{p,r}$ are dense in $\mathcal{H}_F$ [80]. Here $\Delta = \delta^{ab}\partial_a\partial_b$ denotes the Laplacian and we have taken a loop p rather than an open path so that $X_{p,r}$ is transversal. Also unsmeared electric fields are represented through the Fock state ω_F by

$$\omega_F(\hat{A}_r(p)\hat{E}^a(x)\hat{A}_r(p')) = -\frac{\alpha}{2}[X^a_{p,r}(x) - X^a_{p',r}(x)]\omega_F(\hat{h}_{p\circ p',r}) \tag{166}$$

and any other expectation value follows from these and the commutation relations.

Since ω_F defines a positive linear functional we may define a new representation of the algebra $A(p), E^a_r$ by

$$\omega_r(\hat{A}(p)) := \omega_F(\hat{A}_r(p)) \quad \text{and} \quad \omega_r(\hat{A}(p)\hat{E}^a_r(x)\hat{A}(p')) := \omega_F(\hat{A}_r(p)\hat{E}^a(x)\hat{A}_r(p')) \tag{167}$$

called the r-Fock representation. In other words, we have $\omega_r = \omega_F \circ \mathcal{I}_r$.

Since ω_r is a positive linear functional on $C(\overline{\mathcal{A}})$ by construction there exists is a measure μ_r on $\overline{\mathcal{A}}$ that represents ω_r in the sense of the Riesz representation theorem (recall 62). In [81] Velhinho showed that the one-parameter family of measures μ_r are expectedly mutually singular with respect to each other and with respect to the uniform measure μ_0 (that is, the support of one measure is a measure zero set with respect to the other and vice versa).

Result 1: There is a unitary transformation between any of the Hilbert spaces $\mathcal{H}_r$ and their images under $\mathcal{I}_r$ in the usual Fock space $\mathcal{H}_F$. Since finite linear combinations of the $A_r(p)$ for fixed r are still dense in $\mathcal{H}_F$ [80], there exists indeed a polymer like description of the usual n-photon states.

Recall that the Fock vacuum Ω_F is defined to be the zero eigenvalue coherent state, that is, it is annihilated by the annihilation operators

$$\hat{a}(f) := \frac{1}{\sqrt{2\alpha}} \int d^3x f^a \left[\sqrt[4]{-\Delta}\hat{A}_a - i(\sqrt[4]{-\Delta})^{-1}\hat{E}^a\right], \tag{168}$$

where f^a is any transversal smearing field. We then have in fact that $\omega_F(.) = <\Omega_F, .\Omega_F>_{\mathcal{H}_F}$. (For readers familiar with C^*-algebras this means that Ω_F is the cyclic vector that is determined by ω_F through the GNS construction.) The idea is now the following: From (167) we see that we can easily answer any question in the r-Fock representation which has a preimage in the Fock representation, we just have to replace everywhere $A_r(p), E^a(x)$ by $A(p), E^a_r(x)$. Since in the r-Fock representations only exponentials of connections are defined, we should exponentiate the annihilation operators and select the Fock vacuum through the condition

$$e^{i\hat{a}(f)}\Omega_F = \Omega_F. \tag{169}$$

In particular, choosing $f = \sqrt{2\alpha}(\sqrt[4]{-\Delta})^{-1}X_{p,r}$ for some loop p we get

$$e^{\int d^3x X^a_{p,r}[i\hat{A}_a + (\sqrt{-\Delta})^{-1}\hat{E}^a]}\Omega_F = \Omega_F. \tag{170}$$

Using the commutation relations and the Baker–Campell–Hausdorff formula one can write (170) in terms of $\hat{A}_r(p)$ and the exponential of the electric field appearing in (170) times a numerical factor. The resulting expression can then be translated into the r-Fock representation. Denoting the translated expression by $\mathcal{I}_r^{-1}(e^{i\hat{a}(f)})$ we now ask the question, whether there exists a state $\Omega_r \in \mathcal{H}_{\text{kin}} = L_2(\overline{\mathcal{A}}, d\mu_0)$ such that $\mathcal{I}_r^{-1}(e^{i\hat{a}(f)})\Omega_r = \Omega_r$. Remarkably, expanding Ω_r into the charge network basis introduced in Sect. 3.1 one finds a (up to a multiplicative constant) unique solution given by

$$\Omega_r = \sum_c e^{-\frac{\alpha}{2}\sum_{e,e' \in E(\gamma(c))} G^r_{e,e'} n_e(c) n_{e'}(c)} \overline{T_c}, \tag{171}$$

where $c = (\gamma(c), \{n_e(c)\}_{e \in E(\gamma(c))})$ denotes a charge network (the $U(1)$ analogue of a spin network) and

$$G^r_{e,e'} = \int d^3x X^a_{e,r} \sqrt{-\Delta}^{-1} X^b_{e',r} \delta^T_{ab} \,, \tag{172}$$

where $\delta^T_{ab} = \delta_{ab} - \partial_a \Delta^{-1} \partial_b$ denotes the transverse projector.

Exercise 30.
Fill in the gaps that lead from (170) to (172).

Let us discuss this result. First of all, (171) is not normalizable with respect to the inner product on $\mathcal{H}_{\text{kin}}$ and neither are the images of n-photon states or coherent states from $\mathcal{H}_F$. This seems to indicate that the space $\mathcal{H}_{\text{kin}}$ does not play any role for physically interesting states. However, in [74] it was shown that this is not the case: It turns out, that, given a suitable regularization, that one can indeed obtain the expectation values such as $\omega_r(A(p))$ from the formal expression

$$\omega_r(A(p)) := \frac{< \Omega_r, A(p)\Omega_r >}{||\Omega_r||^2} \,, \tag{173}$$

where both numerator and denominator are infinite but the fraction is finite.

Result 2: The polymer images of photon states can be obtained as certain limits of states from $\mathcal{H}_{\text{kin}}$ which therefore is a valid starting point in order to obtain physically interesting representations.

Moreover, as can be expected from the similarity between the formulas (172) and (149) (for $A^{\mathbb{C}} = 0$ corresponding to vacuum $E = A = 0$ in the present case), the states Ω_r also arise from a complexifier, given in this case by

$$C = \frac{1}{2q^2} \int_{\mathbb{R}^3} d^3x [E^a_r \sqrt{-\Delta}^{-1} E^b_r] \delta_{ab} \,. \tag{174}$$

Result 3: The complexifier framework is also able to derive images of n-photon states and usual Fock coherent states from the universal input of a complexifier.

We conclude that at least for the linearized theory the question posed at the beginning of this section could be answered affirmatively: There is indeed a precise framework available for how to accommodate graviton states into the framework of loop quantum gravity. This is a promising result and should have an analog in the full theory.

4 Selection of Open Research Problems

Let us summarize the most important open research problems that have come up during the discussion in these lectures.

i) *Hamiltonian Constraint and Semiclassical States*
The unsettled correctness of the quantum dynamics is the major roadblock to completing the quantization programme of QGR. In order to make progress a better understanding of the kinematical semiclassical sector of the theory is necessary.

ii) *Physical Inner Product*
Even if we had the correct Hamiltonian constraint and the complete space of solutions, at the moment there is no really good idea available of how to construct a corresponding physical inner product because the constraint algebra is not a Lie algebra but an open algebra in the BRST sense so that techniques from rigged Hilbert spaces are not available. A framework for such open algebras must be developed so that an inner product can be constructed at least in principle.

iii) *Dirac Observables*
Not even in classical general relativity do we know enough Dirac observables. For QGR they are mandatory for instance in order to select an inner product by adjointness conditions and in order to arrive at an interpretation of the final theory. A framework of how to define Dirac observables, at least in principle, even at the classical level, would be an extremely important contribution.

iv) *Covariant Formulation*
The connection between the Hamiltonian and the Spin Foam formulation is poorly understood. Without such a connection e.g. a proof of covariance of the canonical formulation on the one hand and a proof for the correct classical limit of the spin foam formulation on the other cannot be obtained using the respective other formulation. One should prove a rigorous Feynman–Kac like formula that allows to switch between these complementary descriptions.

v) *QFT on CST's and Hawking Effect from First Principles*
The low energy limit of the theory in connection with the the construction of semiclassical states must be better understood. Once this is done, fundamental issues such as whether the Hawking effect is merely an artefact of an invalid description by QFT's on CST's while a quantum theory of gravity should be used or whether it is a robust result can be answered. Similar remarks apply to the information paradoxon associated with black holes etc.

vi) *Combinatorial Formulation of the Theory*
The description of a theory in terms of smooth and even analytic structures curves, surfaces etc. at all scales in which the spectra of geometrical operators are discrete at Planck scales is awkward and cannot be the most adequate language. There should be a purely combinatorial formulation in which notions such as topology, differential structure etc. can only have a semiclassical meaning.

vii) *Avoidance of Classical and UV Singularities*
That certain classical singularities are absent in loop quantum cosmology and that certain operators come out finite in the full theory while in the usual perturbative formulation they would suffer from UV singularities are promising results, but they must be better understood. If one could make

contact with perturbative formulations and pin-point exactly why in QGR the usual perturbative UV singularities are absent then the theory would gain a lot more respect in other communities of high energy physicists. There must be some analog of the renormalization group and the running of coupling constants that one usually finds in QFT's and CST's. Similar remarks apply to the generalization of the loop quantum cosmology result to the full theory.

viii) *Contact with String (M)-Theory*
If there is any valid perturbative description of quantum gravity then it is almost certainly string theory. It is conceivable that both string theory and loop quantum gravity are complementary descriptions but by themselves incomplete and that only a fusion of both can reach the status of a fundamental theory. To explore these possibilities, Smolin has launched an ambitious programme [82] which to our mind so far did not raise the interest that it deserves[18]. The contact arises through Chern–Simons theory which is part of both Loop Quantum Gravity and M-Theory [83] (when considered as the high energy limit of 11 dimensional Supergravity). Another obvious starting point is the definition of M-Theory as the quantum supermembrane in 11 dimensions [84], a theory that could be obtained as the quantization of the classical supermembrane by our non-perturbative methods. Finally, a maybe even more obvious connection could be found through the so-called **Pohlmeyer String** [85] which appears to be a method to quantize the string non-perturbatively, without supersymmetry, anomalies or extra dimensions, by working *directly at the level of Dirac observables* which are indeed possible to construct explicitly in this case.

We hope to have convinced the reader that Loop Quantum Gravity is an active and lively approach to a quantum theory of gravity which has produced already many non-trivial results and will continue to do so in the future. There are still a huge number of hard but fascinating problems to be solved of which the above list is at most the tip of an iceberg. If at least a tiny fraction of the readers would decide to dive into this challenging area and help in this endeavour, then these lectures would have been successful.

Acknowledgements

We thank the Heraeus-Stiftung and the organizers, Domenico Giulini, Claus Kiefer and Claus Lämmerzahl, for making this wonderful and successful meeting possible and the participants for creating a stimulating atmosphere through long and deep discussions, very often until early in the morning in the "Bürgerkeller".

[18] That we did not devote a section to this topic in this review is due to the fact that we would need to include an introduction to M-Theory into these lectures which would require too much space. The interested reader is referred to the literature cited.

References

1. T. Thiemann: Introduction to Modern Canonical Quantum General Relativity. gr-qc/0110034
2. C. Rovelli: Loop Quantum Gravity. *Living Reviews* **1**, Pub.-No. 1998-1 (1998) [gr-qc/9710008];
   C. Rovelli: Strings, loops and others: A critical survey of the present approaches to quantum gravity. In N. Dadhich and J.V. Narlikar, editors, *Gravitation and Relativity: at the turn of the Millenium, Proceedings of the 15th International Conference on General Relativty and Gravitation*, (Inter-University Centre for Astrnomy and Astrophysics, Pune, 1998) gr-qc/9803024;
   C. Rovelli: Notes for a Brief History of Quantum Gravity. gr-qc/0006061;
   M. Gaul and C. Rovelli: Loop Quantum Gravity and the Meaning of Diffeomorphism Invariance. *Lect. Notes Phys.* **541**, 277 (2000) [gr-qc/9910079];
   G. Horowitz: Quantum Gravity at the Turn of the Millenium. gr-qc/0011089;
   S. Carlip: Quantum Gravity: A Progress Report. *Rep. Prog. Phys.* **64**, 885 (2001) [gr-qc/0108040];
   A. Ashtekar: Quantum Mechanics of Geometry. gr-qc/9901023
3. E.E. Flanagan and R.M. Wald: Does backreaction enforce the averaged null energy condition in semiclassical gravity? *Phys. Rev. D* **54**, 6233 (1996) [gr-qc/9602052]
4. R.M. Wald: *General Relativity* (The University of Chicago Press, Chicago, 1989)
5. W. Junker: Hadamard states, adiabatic vacua and the construction of physical states for scalar quantum fields on curved spacetime. *Rev. Math. Phys.* **8**, 1091 (1996)
6. M.H. Goroff and A. Sagnotti: *Phys. Lett. B* **160**, 81 (1985);
   M.H. Goroff and A. Sagnotti: *Nucl. Phys. B* **266**, 709 (1986)
7. S. Deser: Non-renormalizability of (last hope) $D = 11$ supergravity with a survey of divergences in quantum gravities. hep-th/9905017
8. O. Lauscher and M. Reuter: Towards nonperturbative renormalizability of quantum Einstein gravity. *Int. J. Mod. Phys. A* **17**, 993 (2002) [hep-th/0112089];
   O. Lauscher and M. Reuter: Is quantum gravity nonperturbatively renormalizable?. *Class. Quant. Grav.* **19**, 483 (2002) [hep-th/0110021]
9. J. Polchinsky: *String Theory*, Vol. 1 and 2 (Cambridge University Press, Cambridge, 1998)
10. D. Marolf and C. Rovelli: Relativistic quantum measurement. *Phys. Rev. D* **66**, 023510 (2002) [gr-qc/0203056]
11. J.B. Hartle and S. W. Hawking: *Phys. Rev. D* **28**, 2960 (1983)
12. J. Ambjorn, J. Jurkiewicz, and R. Loll: Lorentzian and Euclidean quantum gravity: Analytical and numerical results. hep-th/0001124;
    J. Ambjorn, J. Jurkiewicz, and R. Loll: A Non-Perturbative Lorentzian Path Integral for Gravity. *Phys. Rev. Lett.* **85**, 924 (2000) [hep-th/0002050]
    J. Ambjorn, A. Dasgupta, J. Jurkiewicz, and R. Loll: A Lorentzian Cure For Euclidean Troubles. *Nucl. Phys. Proc. Suppl.* **106**, 977 (2002) [hep-th/0201104];
    R. Loll and A. Dasgupta: A proper time cure for the conformal sickness in quantum gravity. *Nucl. Phys. B* **606**, 357 (2001) [hep-th/0103186]
13. M.J. Gotay, J. Isenberg, and J.E. Marsden: Momentum maps and classical relativistic fields. Part 1: Covariant field theory. physics/9801019;
    H.A. Kastrup: Canonical theories of dynamical systems in physics. *Phys. Rept.* **101**, 1 (1983);

I.V. Kanatchikov: Canonical structure of classical field theory in the polymomentum phase space. *Rept. Math. Phys.* **41**, 49 (1998) [hep-th/9709229]
I.V. Kanatchikov: On the field theoretic generalizations of a Poisson algebra. *Rept. Math. Phys.* **40**, 225 (1997) [hep-th/9710069];
C. Rovelli: A Note on the foundation of relativistic mechanics. I: Relativistic observables and relativistic states. [gr-qc/0111037];
C. Rovelli: A Note on the foundation of relativistic mechanics. II: Covariant Hamiltonian general relativity. [gr-qc/0202079]
C. Rovelli: Covariant Hamiltonian formalism for field theory: Symplectic structure and Hamilton-Jacobi equation on the space G. [gr-qc/0207043]

14. A. Ashtekar, L. Bombelli, and O. Reula: The covariant phase space of asymptotically flat gravitational fields. In M. Francaviglia, D. Holm, editors, *Analysis, Geometry and Mechanics: 200 Years After Lagrange* (North-Holland, Amsterdam, 1991)
15. R.E. Peierls: *Proc. R. Soc. Lond. A* **214**, 143 (1952);
B.S. DeWitt: *Dynamical Theory of Groups and Fields* (Gordon & Breach, New York, 1965)
16. C.J. Isham and N. Linden: Quantum temporal logic and decoherence functionals in the histories approach to generalized quantum theory. *J. Math. Phys.* **35**, 5452 (1994) [gr-qc/9405029];
C.J. Isham and N. Linden: Continuous histories and the history group in generalized quantum theory. *J. Math. Phys.* **36**, 5392 (1995) [gr-qc/9503063];
C.J. Isham, N. Linden, K. Savvidou, and S. Schreckenberg: Continuous time and consistent histories. *J. Math. Phys.* **39**, 1818 (1998) [quant-ph/9711031]
17. R.B. Griffiths: *J. Stat. Phys.* **36**, 219 (1984);
R.B. Griffiths: *Found. Phys.* **23**, 1601 (1993);
R. Omnés: *J. Stat. Phys.* **53**, 893 (1988);
R. Omnés: *J. Stat. Phys.* **53**, 933 (1988);
R. Omnés: *J. Stat. Phys.* **53**, 957 (1988);
R. Omnés: *J. Stat. Phys.* **57**, 357 (1989);
R. Omnés: *J. Stat. Phys. Rev. Mod. Phys.* **64**, 339 (1992);
M. Gell-Mann and J.B. Hartle: Classical equations for quantum systems. *Phys. Rev. D* **47**, 3345 (1993) [gr-qc/9210010];
M. Gell-Mann and J.B. Hartle: Equivalent sets of histories and multiple quasiclassical domains. [gr-qc/9404013];
M. Gell-Mann and J.B. Hartle: Strong decoherence. [gr-qc/9509054];
J.B. Hartle: Space-time quantum mechanics and the quantum mechanics of space-time. In B. Julia, J. Zinn-Justin, editors, *Gravitation and quantizations*, Les Houches Session LVII (North-Holland, Amtsredam 1992), page 285 [gr-qc/9304006];
F. Dowker and A. Kent: Properties of consistent histories. *Phys. Rev. Lett.* **75**, 3038 (1995) [gr-qc/9409037];
F. Dowker and A. Kent: On the consistent histories approach to quantum mechanics. *J. Statist. Phys.* **82**, 1575 (1996) [gr-qc/9412067]
18. I.V. Kanatchikov: Precanonical perspective in quantum gravity. *Nucl. Phys. Proc. Suppl.* **88**, 326 (2000) [gr-qc/0004066];
I.V. Kanatchikov: Precanonical quantization and the Schroedinger wave functional. *Phys. Lett. A* **283**, 25 (2001) [hep-th/0012084];
I.V. Kanatchikov: Precanonical quantum gravity: Quantization without the space-time decomposition. *Int. J. Theor. Phys.* **40**, 1121 (2001) [gr-qc/0012074]

19. C.J. Isham and K.N. Savvidou: Quantizing the foliation in history quantum field theory. [quant-ph/0110161]
20. P.A.M. Dirac: *Lectures on Quantum Mechanics*, Belfer Graduate School of Science (Yeshiva University Press, New York, 1964)
21. D.M. Gitman and I.V. Tyutin: *Quantization of Fields with Constraints* (Springer-Verlag, Berlin, 1990);
M. Henneaux, C. Teitelboim: *Quantization of Gauge Systems* (Princeton University Press, Princeton, 1992)
22. P.G. Bergmann and A. Komar: The phase space formulation of general relativity and approaches towards its canonical quantization. *Gen. Rel. Grav.* **1**, 227 (1981);
A. Komar: General relativistic observables via Hamilton–Jacobi functionals. *Phys. Rev. D* **4**, 923 (1971);
A. Komar: Commutator algebra of general relativistic observables. *Phys. Rev. D* **9**, 885 (1974);
A. Komar: Generalized constraint structure for gravitation theory. *Phys. Rev. D* **27**, 2277 (1983);
A. Komar: Consistent factor ordering of general relativistic constraints. *Phys. Rev. D* **20**, 830 (1979);
P.G. Bergmann and A. Komar: The coordinate group symmetries of general relativity. *Int. J. Theor. Phys.* **5**, 15 (1972)
23. S.A. Hojman, K. Kuchar, and C. Teitelboim: Geometrodynamics regained. *Ann. Phys.* (NY) **96**, 88 (1976)
24. C.G. Torre and I.M. Anderson: Symmetries of the Einstein equations. *Phys. Rev. Lett.* **70**, 3525 (1993) [gr-qc/9302033];
C.G. Torre and I.M. Anderson: Classification of generalized symmetries for the vacuum Einstein equations. *Commun. Math. Phys.* **176**, 479 (1996) [gr-qc/9404030]
25. A. Ashtekar: *Phys. Rev. Lett.* **57**, 2244 (1986);
A. Ashtekar: *Phys. Rev. D* **36**, 1587 (1987)
26. G. Immirzi: *Nucl. Phys. Proc. Suppl.* **57**, 65 (1997)
27. F. Barbero: *Phys. Rev. D* **51**, 5507 (1995);
F. Barbero: *Phys. Rev. D* **51**, 5498 (1995)
28. M. Nakahara: *Geometry, Topology and Physics* (Institute of Physics Publishing, Bristol, 1998)
29. D. Giulini and D. Marolf: On the generality of refined algebraic quantization. *Class. Quant. Grav.* **16**, 2479 (1999) [gr-qc/9812024]
30. N.M.J. Woodhouse: *Geometric Quantization*, 2nd. edition (Clarendon Press, Oxford, 1991)
31. R. Gambini, A. Trias: *Phys. Rev. D* **22**, 1380 (1980);
C. Di Bartolo, F. Nori, R. Gambini, and A. Trias: *Lett. Nuov. Cim.* **38**, 497 (1983);
R. Gambini, A. Trias: *Nucl. Phys. B* **278**, 436 (1986)
32. M.F. Atiyah: Topological Quantum Field Theories. *Publ. Math. IHES* **68**, 175 (1989);
M.F. Atiyah: *The Geometry of Physics and Knots* (Cambridge University Press, Cambridge, 1990)
33. T. Jacobson and L. Smolin: Nonperturbative quantum geometries. *Nucl. Phys. B* **299**, 295 (1988);
C. Rovelli and L. Smolin: Loop space representation of quantum general relativity. *Nucl. Phys. B* **331**, 80 (1990)
34. R. Giles: *Phys. Rev. D* **8** (1981) 2160

35. H. Sahlmann: When do measures on the space of connections support the triad operators of loop quantum gravity. [gr-qc/0207112];
H. Sahlmann: Some comments on the representation theory of the algebra underlying loop quantum gravity. [gr-qc/0207111];
H. Sahlmann and T. Thiemann: Representation theory of diffeomorphism invariant gauge theories. in preparation
36. A. Ashtekar and J. Lewandowski: *J. Math. Phys.* **36**, 2170 (1995)
37. J. Velhinho: A groupoid approach to spaces of generalized connections. hep-th/0011200
38. A. Ashtekar and C.J. Isham: *Class. Quantum Grav.* **9**, 1433 (1992)
39. A. Ashtekar and J. Lewandowski: Representation theory of analytic holonomy $C^\star$ algebras'. In J. Baez, editor, *Knots and Quantum Gravity*, (Oxford University Press, Oxford, 1994)
40. C. Rovelli and L. Smolin: Spin networks and quantum gravity. *Phys. Rev. D* **53**, 5743 (1995);
J. Baez: Spin Networks in non-perturbative quantum gravity. In L. Kauffman, editor, *The Interface of Knots and Physics*, (American Mathematical Society, Providence, Rhode Island, 1996) [gr-qc/9504036]
41. N.J. Vilenkin: *Special Functions and the Theory of Group Representations* (American Mathematical Society, Providence, Rhode Island, 1968)
42. T. Thiemann and O. Winkler: Gauge field theory coherent states (GCS): IV. Infinite tensor product and thermodynamical limit. *Class. Quantum Grav.* **18**, 4997 (2001) [hep-th/0005235];
M. Arnsdorf. Loop quantum gravity on noncompact spaces: *Nucl. Phys. B* **577**, 529 (2000) [gr-qc/9909053]
43. A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão, and T. Thiemann. Quantization for diffeomorphism invariant theories of connections with local degrees of freedom. *J. Math. Phys.* **36**, 6456 (1995) [gr-qc/9504018]
44. C. Rovelli and L. Smolin: Discreteness of volume and area in quantum gravity. *Nucl. Phys. B* **442**, 593 (1995) (Erratum: *Nucl. Phys. B* **456**, 734 (1995));
A. Ashtekar and J. Lewandowski: Quantum theory of geometry I: Area operators. *Class. Quantum Grav.* **14**, A55 (1997)
45. A. Ashtekar and J. Lewandowski: Quantum theory of geometry II: Volume operators. *Adv. Theo. Math. Phys.* **1**, 388 (1997);
J. Lewandowski: *Class. Quantum Grav.* **14**, 71 (1997);
R. Loll: Spectrum of the volume operator in quantum gravity. *Nucl. Phys. B* **460**, 143 (1996) [gr-qc/9511030];
R. De Pietri and C. Rovelli: Geometry eigenvalues and scalar product from recoupling theory in loop quantum gravity. *Phys. Rev. D* **54**, 2664 (1996) [gr-qc/9602023];
T. Thiemann: Closed formula for the matrix elements of the volume operator in canonical quantum gravity. *J. Math. Phys.* **39**, 3347 (1998) [gr-qc/9606091]
46. T. Thiemann: A length operator for canonical quantum gravity. *Journ. Math. Phys.* **39**, 3372 (1998) [gr-qc/9606092]
47. T. Thiemann: Spatially diffeomorphism invariant geometrical operators in quantum general relativity. Work in progress
48. T. Thiemann: Anomaly-free Formulation of non-perturbative, four-dimensional Lorentzian Quantum Gravity. *Phys. Lett. B* **380**, 257 (1996) [gr-qc/9606088];
T. Thiemann: Quantum spin dynamics (QSD). *Class. Quantum Grav.* **15**, 839 (1998) [gr-qc/9606089];

T. Thiemann: Quantum spin dynamics (QSD): II. The kernel of the Wheeler–DeWitt constraint operator. *Class. Quantum Grav.* **15**, 875 (1998), [gr-qc/9606090];
T. Thiemann: Quantum spin dynamics (QSD): III. Quantum constraint algebra and physical scalar product in quantum general relativity. *Class. Quantum Grav.* **15**, 1207 (1998) [gr-qc/9705017];
T. Thiemann: Quantum spin dynamics (QSD): IV. 2+1 Euclidean quantum gravity as a model to test 3+1 Lorentzian quantum gravity. *Class. Quantum Grav.* **15**, 1249 (1998) [gr-qc/9705018];
T. Thiemann: Quantum spin dynamics (QSD): V. Quantum Gravity as the natural regulator of the Hamiltonian constraint of matter quantum field theories. *Class. Quantum Grav.* **15**, 1281 (1998) [gr-qc/9705019];
T. Thiemann: Quantum spin dynamics (QSD): VI. Quantum Poincaré algebra and a quantum positivity of energy theorem for canonical quantum gravity. *Class. Quantum Grav.* **15**, 1463 (1998), [gr-qc/9705020];
T. Thiemann: Kinematical Hilbert spaces for fermionic and Higgs quantum field theories. *Class. Quantum Grav.* **15**, 1487 (1998) [gr-qc/9705021]

49. H. Sahlmann and T. Thiemann: Towards the QFT on curved spacetime limit of QGR. 1. A general scheme. [gr-qc/0207030];
H. Sahlmann and T. Thiemann: Towards the QFT on curved spacetime limit of QGR. 2. A Concrete implementation. [gr-qc/0207031]
50. D. Marolf and J. Lewandowski: Loop Constraints: A habitat and their algebra. *Int. J. Mod. Phys. D* **7**, 299 (1998) [gr-qc/9710016];
R. Gambini, J. Lewandowski, D. Marolf, and J. Pullin: On the consistency of the constraint algebra in spin network gravity. *Int. J. Mod. Phys. D* **7**, 97 (1998) [gr-qc/9710018]
51. M. Gaul and C. Rovelli: A generalized Hamiltonian contraint operator in loop quantum gravity and its simplest Euclidean matrix elements. *Class. Quantum Grav.* **18**, 1593 (2001) [gr-qc/0011106]
52. C. Kiefer: *Lect. Notes Phys.* **541**, 158 (2000) [gr-qc/9906100]; A.O. Barvinsky: Quantum cosmology at the turn of the millenium. [gr-qc/0101046];
J.B. Hartle: Quantum cosmology: Problems for the 21st century. [gr-qc/9701022]
53. M. Bojowald: Loop quantum cosmology. I. Kinematics. *Class. Quantum Grav.* **17**, 1489 (2000) [gr-qc/9910103];
M. Bojowald: Loop quantum cosmology. II. Volume operators. *Class. Quantum Grav.* **17**, 1509 (2000) [gr-qc/9910104];
M. Bojowald: Loop quantum cosmology. III. Wheeler-DeWitt operators. *Class. Quantum Grav.* **18**, 1055 (2001) [gr-qc/0008052];
M. Bojowald: Loop quantum cosmology. IV. Discrete time evolution. *Class. Quantum Grav.* **18**, 1071 (2001) [gr-qc/0008053];
M. Bojowald: Absence of singularity in Loop quantum cosmology. *Phys. Rev. Lett.* **86**, 5227 (2001) [gr-qc/0102069];
M. Bojowald: Dynamical initial conditions in quantum cosmology. *Phys. Rev. Lett.* **87**, 121301 (2001) [gr-qc/0104072];
M. Bojowald: The inverse scale factor in isotropic quantum geometry. *Phys. Rev. D* **64**, 084018 (2001) [gr-qc/0105067];
M. Bojowald:. The semiclassical limit of loop quantum cosmology. *Class. Quantum Grav.* **18**, L109 (2001) [gr-qc/0105113];
M. Bojowald: Inflation from quantum geometry. gr-qc/0206054

54. J.C. Baez: An introduction to spin foam models of quantum gravity and BF Theory. *Lect. Notes Phys.* **543**, 25 (2000) [gr-qc/9905087];
J.C. Baez: Spin foam models. *Class. Quantum Grav.* **15**, 1827 (1998) [gr-qc/9709052];
J.W. Barrett: State sum models for quantum gravity. gr-qc/0010050;
J.W. Barrett: Quantum gravity as topological quantum field theory. *J. Math. Phys.* **36**, 6161 (1995) [gr-qc/9506070];
A. Perez: Spin foam models for quantum gravity. to appear
55. M. Reisenberger and C. Rovelli: Sum over surfaces form of loop quantum gravity. *Phys. Rev. D* **56**, 3490 (1997)
56. G. Roepstorff: *Path Integral Approach to Quantum Physics:An Introduction* (Springer Verlag, Berlin, 1994)
57. S. Holst: Barbero's Hamiltonian derived from a generalized Hilbert-Palatini action. *Phys. Rev. D* **53**, 5966 (1996) [gr-qc/9511026];
N. Barros e Sá: Hamiltonian analysis of general relativity with the Immirzi parameter. *Int. J. Mod. Phys. D* **10**, 261 (2001) [gr-qc/0006013]
58. L. Freidel, K. Krasnov, and R. Puzio: BF description of higher dimensional gravity theories. *Adv. Theor. Math. Phys.* **3**, 1289 (1999) [hep-th/9901069]
59. J.W. Barrett, L. Crane: "Relativistic spin networks and quantum gravity". *J. Math. Phys.* **39**, 3296 (1998) [gr-qc/9709028]
60. J.C. Baez and J.W. Barrett: Integrability of relativistic spin networks. gr-qc/0101107;
A. Perez and C. Rovelli: Spin foam model for Lorentzian general relativity. *Phys. Rev. D* **63**, (2001) 041501, [gr-qc/0009021];
L. Crane, A. Perez, and C. Rovelli: A finiteness proof for the Lorentzian state sum spin foam model for quantum general relativity. gr-qc/0104057;
L. Crane, A. Perez, and C. Rovelli: Perturbative finiteness in spin-foam quantum gravity. *Phys. Rev. Lett.* **87**, 181301 (2001)
61. D.V. Boulatov: *Mod. Phys. Lett. A* **7**, 1629 (1992);
H. Ooguri: *Mod. Phys. Lett. A* **7**, 2799 (1992)
62. F. Markopoulou and L. Smolin: Causal evolution of spin networks. *Nucl. Phys. B* **508**, 409 (1997) [gr-qc/9702025];
F. Markopoulou: Dual formulation of spin network evolution. gr-qc/9704013;
F. Markopoulou and L. Smolin: Quantum geometry with intrinsic local causality. *Phys. Rev. D* **58**, 084032 (1998) [gr-qc/9712067];
F. Markopoulou: The internal description of a causal set: What the universe like from inside. *Commun. Math. Phys.* **211**, 559 (2000) [gr-qc/9811053];
F. Markopoulou: Quantum causal histories. *Class. Quant. Grav.* **17**, 2059 (2000) [hep-th/9904009];
F. Markopoulou: An insider's guide to quantum causal histories. *Nucl. Phys. Proc. Suppl.* **88**, 308 (2000) [hep-th/9912137]
63. F. Markopoulou: An algebraic approach to coarse graining. hep-th/0006199
64. J.C. Baez, J.D. Christensen, T.R. Halford, and D.C. Tsang: Spin foam models of Riemannian quantum gravity. *Class. Quant. Grav.* **19**, 4627 (2002) [gr-qc/0202017];
A. Perez: Spin foam quantization of Plebanski's action. *Adv. Theor. Math. Phys.* **5** 947 (2002) [gr-qc/0203058]
65. M. Bojowald and A. Perez: Spin foam quantization and anomalies. In preparation

66. J.D. Bekenstein: Black holes and entropy. *Phys. Rev. D* **7**, 2333 (1973);
J.D. Bekenstein: Generalized second law for thermodynamics in black hole physics. *Phys. Rev. D* **9**, 3292 (1974);
S.W. Hawking: Particle creation by black holes. *Commun. Math. Phys.* **43**, 199 (1975)
67. K. Krasnov: On statistical mechanics of gravitational systems. *Gen. Rel. Grav.* **30**, 53 (1998) [gr-qc/9605047];
C. Rovelli: Black hole entropy from loop quantum gravity. *Phys. Rev. Lett.* **77**, 3288 (1996) [gr-qc/9603063]
68. A. Ashtekar, C. Beetle, O. Dreyer, S. Fairhurst, B. Krishnan, J. Lewandowski, and J. Wisniewski: Isolated horizons and their applications. *Phys. Rev. Lett.* **85**, 3564 (2000) [gr-qc/0006006];
A. Ashtekar: Classical and quantum physics of isolated horizons. *Lect. Notes Phys.* **541**, 50 (2000);
A. Ashtekar: Interface of general relativity, quantum physics and statistical mechanics: Some recent developments. *Ann. Phys.* **9**, 178 (2000) [gr-qc/9910101]
69. A. Ashtekar, A. Corichi, and K. Krasnov: Isolated horizons: The classical phase space. *Adv. Theor. Math. Phys.* **3**, 419 (2000) [gr-qc/9905089];
A. Ashtekar, J. C. Baez, and K. Krasnov: Quantum geometry of isolated horizons and black hole entropy. *Adv. Theor. Math. Phys.* **4**, 1 (2001) [gr-qc/0005126]
70. T. Regge and C. Teitelboim: Role of surface integrals in the Hamiltonian formulation of general relativity. *Ann. Phys.* **88**, 286 (1974)
71. L. Smolin: Linking topological quantum field theory and non-perturbative quantum gravity". *J. Math. Phys.* **36**, 6417 (1995) [gr-qc/9505028]
72. S. Axelrod, S.D. Pietra, and E. Witten: Geometric quantization of Chern-Simons gauge theory. *J. Diff. Geom.* **33**, 787 (1991)
73. G. 't Hooft: The Holographic Principle: Opening Lecture. In *Basics and Highlights in Fundamental Physics*, Opening Lecture Erice 1999. [hep-th/0003004]
74. T. Thiemann: Reality conditions inducing transforms for quantum gauge field theories and quantum gravity. *Class. Quantum Grav.* **13**, 1383 (1996) [gr-qc/9511057];
T. Thiemann: An account of transforms on $\mathcal{A}/\mathcal{G}$. *Acta Cosmologica* **21**, 145 (1995) [gr-qc/9511049];
T. Thiemann: Gauge field theory coherent states (GCS): I. General properties. *Class. Quant. Grav.* **18** 2025 (2001) [hep-th/0005233];
T. Thiemann: Complexifier coherent states for quantum general relativity. [gr-qc/0206037]
75. T. Thiemann: Quantum spin dynamics (QSD): VII. Symplectic structures and continuum lattice formulations of gauge field theories. *Class. Quant. Grav.* **18**, 3293 (2001) [hep-th/0005232];
T. Thiemann and O. Winkler: Gauge field theory coherent states (GCS): II. Peakedness properties. *Class. Quant. Grav.* **18**, 2561 (2001) [hep-th/0005237];
T. Thiemann and O. Winkler: Gauge field theory coherent states (GCS): III. Ehrenfest theorems. *Class. Quantum Grav.* **18** (2001) 4629-4681, [hep-th/0005234];
H. Sahlmann, T. Thiemann, and O. Winkler: Coherent states for canonical quantum general relativity and the Infinite tensor product extension. *Nucl. Phys. B* **606**, 401 (2001) [gr-qc/0102038]
76. B.C. Hall: *Journ. Funct. Analysis* **122**, 103 (1994);
B.C. Hall and J.J. Mitchell: Coherent states on spheres. *J. Math. Phys.* **43**, 1211 (2002) [quant-ph/0109086];

B.C. Hall and J.J. Mitchell: The large radius limit for coherent states on spheres. quant-ph/0203142
77. G. Amelino-Camelia: Are we at dawn with quantum gravity phenomenology? *Lect. Notes Phys.* **541**, 1 (2000) [gr-qc/9910089];
G. Amelino-Camelia, John R. Ellis, N.E. Mavromatos, D.V. Nanopoulos, and S. Sarkar: Potential sensitivity of gamma ray burster observations to wave dispersion in vacuo. *Nature* **393**, 763 (1998) [astro-ph/9712103];
R. Gambini and J. Pullin: Nonstandard optics from quantum spacetime. *Phys. Rev. D* **59** 124021 (1999) [gr-qc/9809038];
R. Gambini and J. Pullin: Quantum gravity experimental physics? *Gen. Rel. Grav.* **31**, 1631 (1999);
J. Alfaro, H.A. Morales-Tecotl, and L.F. Urrutia: Quantum gravity corrections to neutrino propagation. *Phys. Rev. Lett.* **84**, 2318 (2000) [gr-qc/9909079];
J. Alfaro, H.A. Morales-Tecotl, and L.F. Urrutia: Loop quantum gravity and light propagation. *Phys. Rev. D* **65**, 103509 (2002) [hep-th/0108061]
78. S.D. Biller et al.: *Phys. Rev. Lett.* **83**, 2108 (1999)
79. A. Ashtekar, C. Rovelli, and L. Smolin: Gravitons and loops. *Phys. Rev. D* **44**, 1740 (1991) [hep-th/9202054]
80. M. Varadarajan: Fock representations from $U(1)$ holonomy algebras. *Phys. Rev. D* **61**, 104001 (2000) [gr-qc/0001050];
M. Varadarajan: Photons from quantized electric flux representations. *Phys. Rev. D* **64**, 104003 (2001) [gr-qc/0104051];
M. Varadarajan: Gravitons from a loop representation of linearized gravity. *Phys. Rev. D* **66**, 024017 (2002) [gr-qc/0204067]
81. J. Velhinho: Invariance properties of induced Fock measures for $U(1)$ holonomies. *Commun. Math. Phys.* **227**, 541 (2002) [math-ph/0107002]
82. L. Smolin: Strings as Perturbations of Evolving Spin Networks. *Nucl. Phys. Proc. Suppl.* **88**, 103 (2000) [hep-th/9801022]¿;
L. Smolin: A holographic formulation of quantum general relativity. *Phys. Rev. D* **61**, 084007 (2000) [hep-th/9808191];
L. Smolin: Towards a background independent approach to M theory. hep-th/9808192;
L. Smolin: The cubic matrix model and duality between strings and loops. [hep-th/0006137];
L. Smolin: A candidate for a background independent formulation of M theory. *Phys. Rev. D* **62**, 086001 (2000) [hep-th/9903166];
L. Smolin: The exceptional Jordan algebra and the matrix string. hep-th/0104050;
Y. Ling and L. Smolin: Eleven-dimensional supergravity as a constrained topological field theory. *Nucl. Phys. B* **601**, 191 (2001) [hep-th/0003285];
Y. Ling and L. Smolin: Supersymmetric Spin networks and quantum supergravity. *Phys. Rev. D* **61**, 044008 (2000) [hep-th/9904016];
Y. Ling and L. Smolin: Holographic formulation of quantum supergravity. *Phys. Rev. D* **63**, 064010 (2001) [hep-th/0009018]
83. L. Smolin: M Theory as a matrix extension of Chern-Simons theory. *Nucl. Phys. B* **591** 227 (2000) [hep-th/0002009];
L. Smolin: Quantum gravity with a positive cosmological constant. hep-th/0209079
84. R. Helling and H. Nicolai. Supermebranes and matrix theory. hep-th/9809103
85. K. Pohlmeyer: A group theoretical approach to the quantization of the free relativistic closed string. *Phys. Lett. B* **119**, 100 (1982);

K. Pohlmeyer and K.H. Rehren: Algebraic properties of the invariant charges of the Nambu-Goto theory. *Commun. Math. Phys.* **105**, 593 (1986);
K. Pohlmeyer and K.H. Rehren: The algebra formed by the charges of the Nambu-Goto theory: Identification of a maximal abelean subalgebra. *Commun. Math. Phys.* **114**, 55 (1988);
K. Pohlmeyer and K.H. Rehren: The algebra formed by the charges of the Nambu-Goto theory: Their geometric origin and their completeness. *Commun. Math. Phys.* **114**, 177 (1988);
K. Pohlmeyer: The invariant charges of the Nambu-Goto theory in WKB approximation. *Commun. Math. Phys.* **105**, 629 (1986);
K. Pohlmeyer: The algebra formed by the charges of the Nambu-Goto theory: Casimir elements. *Commun. Math. Phys.* **114**, 351 (1988);
K. Pohlmeyer: Uncovering the detailed structure of the algebra formed by the invariant charges of closed bosonic Strings Moving in (1+2)-Dimensional Minkowski space. *Commun. Math. Phys.* **163**, 629 (1994);
K. Pohlmeyer: The invariant charges of the Nambu-Goto theory: Non-additive composition laws. *Mod. Phys. Lett. A* **10**, 295 (1995);
K. Pohlmeyer: The Nambu-Goto theory of closed bosonic strings moving in (1+3)-dimensional Minkowski space: The quantum algebra of observables. *Ann. Phys.* **8**, 19 (1999) [hep-th/9805057];
K. Pohlmeyer and M. Trunk: The invariant charges of the Nambu-Goto theory: Quantization of non-additive composition laws. hep-th/0206061;
G. Handrich and C. Nowak: The Nambu-Goto theory of closed bosonic strings moving in (1+3)-dimensional Minkowski space: The construction of the quantum algebra of observables up to degree five. *Ann. Phys.* **8**, 51 (1999) [hep-th/9807231];
G. Handrich: Lorentz covariance of the quantum algebra of observables: Nambu-Goto strings in 3+1 dimensions. *Int. J. Mod. Phys. A* **17**, 2331 (2002);
G. Handrich, C. Paufler, J.B. Tausk, and M. Walter: The representation of the algebra of observables of the closed bosonic string in 1+3 dimensions: Calculation to order $\hbar^7$. math-ph/0210024;
C. Meusburger and K.H. Rehren: Algebraic quantization of the closed bosonic string. math-ph/0202041

A Discrete History of the Lorentzian Path Integral

Renate Loll

Institute for Theoretical Physics, Utrecht University, Leuvenlaan 4, 3584 CE Utrecht

Abstract. In these lecture notes, I describe the motivation behind a recent formulation of a non-perturbative gravitational path integral for Lorentzian (instead of the usual Euclidean) space-times, and give a pedagogical introduction to its main features. At the regularized, discrete level this approach solves the problems of (i) having a well-defined Wick rotation, (ii) possessing a coordinate-invariant cutoff, and (iii) leading to *convergent* sums over geometries. Although little is known as yet about the existence and nature of an underlying continuum theory of quantum gravity in four dimensions, there are already a number of beautiful results in $d = 2$ and $d = 3$ where continuum limits have been found. They include an explicit example of the inequivalence of the Euclidean and Lorentzian path integrals, a non-perturbative mechanism for the cancellation of the conformal factor, and the discovery that causality can act as an effective regulator of quantum geometry.

1 Introduction

The desire to understand the quantum physics of the gravitational interactions lies at the root of many recent developments in theoretical high-energy physics. By *quantum gravity* I will mean a consistent fundamental quantum description of space-time geometry (with or without matter) whose classical limit is general relativity. Among the possible ramifications of such a theory are a model for the structure of space-time near the Planck scale, a consistent calculational scheme to compute gravitational effects at all energies, a description of (quantum) geometry near space-time singularities and a non-perturbative quantum description of four-dimensional black holes. It might also help us in understanding cosmological issues about the beginning (and end?) of our universe, although it should be said that some questions (for example, that of the "initial conditions") are likely to remain outside the scope of any physical theory.

From what we know about the quantum dynamics of the other fundamental interactions it seems eminently plausible that also the gravitational excitations should at very short scales be governed by quantum laws, so why have we so far not been able to determine what they are? – One obvious obstacle is the difficulty in finding any direct or indirect evidence for quantum gravitational effects, be they experimental or observational, which could provide a feedback for model-building. A theoretical complication is that the outstanding problems mentioned above require a non-perturbative treatment; it is not sufficient to know the first few terms of a perturbation series. This is true for both conven-

R. Loll, A Discrete History of the Lorentzian Path Integral, Lect. Notes Phys. **631**, 137–171 (2003)
http://www.springerlink.com/

tional perturbative path integral expansions of gravity or supergravity[1] and a perturbative expansion in the string coupling in the case of unified approaches. One avenue to take is to search for a *non-perturbative* definition of such a theory, where the initial input of any fixed "background metric" is inessential (or even undesirable), and where "space–time" is determined *dynamically.* Whether or not such an approach necessarily requires the inclusion of higher dimensions and fundamental supersymmetry is currently unknown. As we will see in the course of these lecture notes, it is perfectly conceivable that one can do without.

Such a non-perturbative viewpoint is very much in line with how one proceeds in classical general relativity, where a metric space-time $(M, g_{\mu\nu})$ (+matter) emerges only as a *solution* to the Einstein equations

$$R_{\mu\nu}[g] - \frac{1}{2} g_{\mu\nu} R[g] + \Lambda g_{\mu\nu} = -8\pi G_N T_{\mu\nu}[\Phi], \tag{1}$$

which define the classical dynamics on the space $\mathcal{M}(M)$, the space of all metrics on a given differentiable manifold M. The analogous question I want to address in the quantum theory is

> Can we obtain "quantum space-time" as a solution to a set of non-perturbative quantum equations of motion on a suitable quantum analogue of $\mathcal{M}(M)$ or rather, of the space of geometries, $\mathrm{Geom}(M) := \mathcal{M}(M)/\mathrm{Diff}(M)$?

This is not a completely straightforward task. Whichever way we want to proceed non-perturbatively, if we give up the privileged role of a flat, Minkowskian background space–time on which the quantization is to take place, we also have to abandon the central role usually played by the Poincaré group, and with it most standard quantum field-theoretic tools for regularization and renormalization. If one works in a continuum metric formulation of gravity, the symmetry group of the Einstein action is instead the group Diff(M) of diffeomorphisms on M, which in terms of local charts are simply the smooth invertible coordinate transformations $x^\mu \mapsto y^\mu(x^\mu)$.[2]

I will in the following describe a particular path integral approach to quantum gravity, which is non-perturbative from the outset in the sense of being defined on the "space of all geometries" (to be defined later), without distinguishing any background metric structure (see also [1,2] for related reviews). This is closely related in spirit with the canonical approach of loop quantum gravity [3] and its more recent incarnations using so-called spin networks [4,5], although there are

[1] Of course, we already know that in these cases a quantization based on a decomposition $g_{\mu\nu}(x) = \eta^{\mathrm{Mink}}_{\mu\nu} + \sqrt{G_N}\ h_{\mu\nu}(x)$, for a linear spin-2 perturbation around Minkowski space leads to a non-renormalizable theory.

[2] One should not get confused here by the fact that in gauge formulations of gravity which work with vierbeins e^a_μ instead of the metric tensor $g_{\mu\nu}$, one has an additional local invariance under SO(3,1)-frame rotations, i.e. elements of the Lorentz group, in addition to diffeomorphism invariance. Nevertheless, this formulation is still not invariant under *global* Lorentz- or Poincaré-transformations.

significant differences in methodology and attitude. "Non-perturbative" means in a covariant context that the path sum or integral will have to be performed explicitly, and not just evaluated around its stationary points, which can only be achieved in an appropriate regularization. The method I will employ uses a discrete lattice regularization as an intermediate step in the construction of the quantum theory. However, unlike in lattice QCD, the lattice and its geometric properties will not be part of a static background structure, but dynamical quantities, as befits a theory of *quantum geometry*.

2 Quantum Gravity from Dynamical Triangulations

In this section I will explain how one may construct a theory of quantum gravity from a non-perturbative path integral, and what logic has led my collaborators and me to consider the method of Lorentzian dynamical triangulations to achieve this. The method is minimal in the sense of employing standard tools from quantum field theory and the theory of critical phenomena and adapting them to the case of *generally covariant systems*, without invoking any symmetries beyond those of the classical theory. At an intermediate stage of the construction, we use a regularization in terms of simplicial "Regge geometries", that is, piecewise linear manifolds. In this approach, "computing the path integral" amounts to a conceptually simple and geometrically transparent "counting of geometries", with additional weight factors which are determined by the Einstein action. This is done first of all at a regularized level. Subsequently, one searches for interesting continuum limits of these discrete models which are possible candidates for theories of quantum gravity, a step that will always involve a renormalization. From the point of view of statistical mechanics, one may think of Lorentzian dynamical triangulations as a new class of statistical models of Lorentzian random surfaces in various dimensions, whose building blocks are flat simplices which carry a "time arrow", and whose dynamics is entirely governed by their intrinsic geometric properties.

Before describing the details of the construction, it may be helpful to recall the path integral representation for a (one-dimensional) non-relativistic particle [6]. The time evolution of the particle's wave function ψ may be described by the integral equation

$$\psi(x'', t'') = \int_{\mathbf{R}} G(x'', x'; t'', t')\psi(x', t'), \tag{2}$$

where the propagator or Feynman kernel G is defined through a limiting procedure,

$$G(x'', x'; t'', t') = \lim_{\epsilon \to 0} A^{-N} \prod_{k=1}^{N-1} \int dx_k \; \mathrm{e}^{i \sum_{j=0}^{N-1} \epsilon L(x_{j+1}, (x_{j+1}-x_j)/\epsilon)}. \tag{3}$$

The time interval $t'' - t'$ has been discretized into N steps of length $\epsilon = (t'' - t')/N$, and the right-hand side of (3) represents an integral over all piecewise

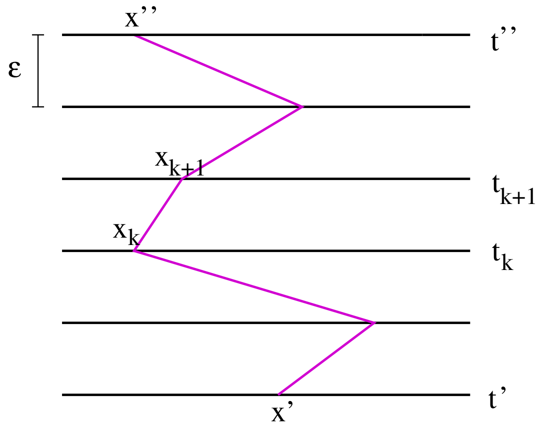


Fig. 1. A piecewise linear particle path contributing to the discrete Feynman propagator

linear paths $x(t)$ of a "virtual" particle propagating from x' to x'', illustrated in Fig. 1.

The prefactor A^{-N} is a normalization and L denotes the Lagrange function of the particle. Knowing the propagator G is tantamount to having solved the quantum dynamics. This is the simplest instance of a *path integral*, and is often written schematically as

$$G(x', t'; x'', t'') = \int \mathcal{D}x(t)\ \mathrm{e}^{iS[x(t)]}, \tag{4}$$

where $\mathcal{D}x(t)$ is a functional measure on the "space of all paths", and the exponential weight depends on the classical action $S[x(t)]$ of a path. Recall also that this procedure can be defined in a mathematically clean way if we Wick-rotate the time variable t to imaginary values $t \mapsto \tau = it$, thereby making all integrals real [7].

Can a similar strategy work for the case of Einstein gravity? As an analogue of the particle's position we can take the geometry $[g_{ij}(x)]$ (ie. an equivalence class of spatial metrics) of a constant-time slice. Can one then define a gravitational propagator

$$G([g'_{ij}], [g''_{ij}]) = \int_{\mathrm{Geom(M)}} \mathcal{D}[g_{\mu\nu}]\ \mathrm{e}^{iS^{\mathrm{Einstein}}[g_{\mu\nu}]} \tag{5}$$

from an initial geometry $[g']$ to a final geometry $[g'']$ (Fig. 2) as a limit of some discrete construction analogous to that of the non-relativistic particle (3)? And crucially, what would be a suitable class of "paths", that is, space–times $[g_{\mu\nu}]$ to sum over?

Setting aside the question of the physical meaning of an expression like (5), gravitational path integrals in the continuum are extremely ill-defined. Clearly, *defining* a fundamental theory of quantum gravity via a perturbation series

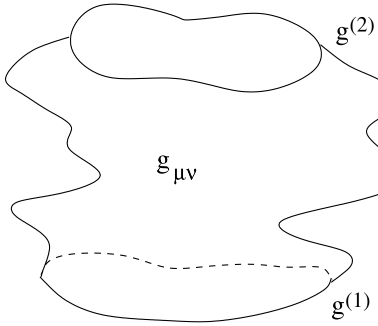


Fig. 2. The time-honoured way [8] of illustrating the gravitational path integral as the propagator from an initial to a final spatial boundary geometry

in the gravitational coupling does not work because of its perturbative non-renormalizability. So, is there a chance we might simply be able to *do* the integration $\int \mathcal{D}[g_{\mu\nu}]$ in a meaningful way? Firstly, there is no obvious way to parametrize "geometries", which means that in practice one always has to start with gauge-*co*variant fields, and gauge-fix. Unfortunately, this gives rise to Faddeev–Popov determinants whose non-perturbative evaluation is exceedingly difficult. A similar problem already applies to the action itself, which is by no means quadratic, no matter what we choose as our basic fields. How then can the integration over $\exp(iS)$ possibly be performed? Part of the problem is clearly also the complex nature of this integrand, with no obvious choice of a Wick rotation in the context of a theory with fluctuating geometric degrees of freedom. Secondly, since we are dealing with a field theory, some kind of regularization will be necessary, and the challenge here is to find a procedure that does not violate diffeomorphism-invariance.

In brief, the strategy I will be following starts from a regularized version of the space Geom(M) of all geometries. A regularized path integral $G(a)$ can be defined which depends on an ultraviolet cutoff a and is *convergent* in a non-trivial region of the space of coupling constants. Taking the continuum limit corresponds to letting $a \to 0$. The resulting continuum theory – if it can be shown to exist – is then investigated with regard to its geometric properties and in particular its semiclassical limit.

3 Brief Summary of Discrete Gravitational Path Integrals

Trying to construct non-perturbative path integrals for gravity from sums over discretized geometries is not a new idea. The approach of *Lorentzian dynamical triangulations* draws from older work in this area, but differs from it in several significant aspects as we shall see in due course.

Inspired by the successes of lattice gauge theory, attempts to describe quantum gravity by similar methods have been popular on and off since the late 70's. Initially the emphasis was on gauge-theoretic, first-order formulations of gravity,

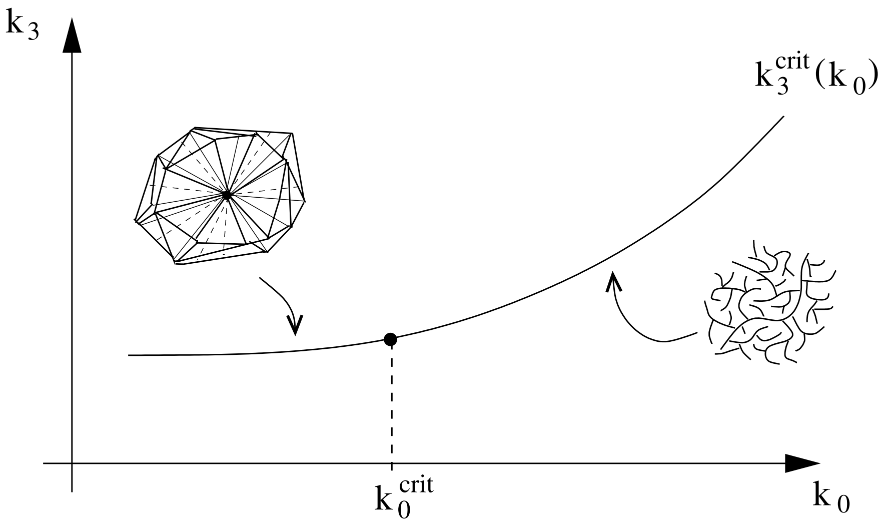


Fig. 3. The phase diagram of three- and four-dimensional Euclidean dynamical triangulations

usually based on (compactified versions of) the Lorentz group, followed in the 80's by "quantum Regge calculus", an attempt to represent the gravitational path integral as an integral over certain piecewise linear geometries (see [9] and references therein), which had first made an appearance in approximate descriptions of *classical* solutions of the Einstein equations. A variant of this approach by the name of "dynamical triangulation(s)" attracted a lot of interest during the 90's, partly because it had proved a powerful tool in describing two-dimensional quantum gravity (see the textbook [10] and lecture notes [11] for more details).

The problem is that none of these attempts have so far come up with convincing evidence for the existence of an underlying continuum theory of *four*-dimensional quantum gravity. This conclusion is drawn largely on the basis of numerical simulations, so it is by no means water-tight, although one can make an argument that the "symptoms" of failure are related in the various approaches [12]. What goes wrong generically seems to be a dominance in the continuum limit of highly degenerate geometries, whose precise form depends on the approach chosen. One would of course expect that non-smooth geometries play a decisive role, in the same way as it can be shown in the particle case that the support of the measure in the continuum limit is on a set of nowhere differentiable paths. However, what seems to happen in the case of the path integral for four-geometries is that the structures obtained are *too* wild, in the sense of not generating, even at coarse-grained scales, an effective geometry whose dimension is anywhere near four.

The schematic phase diagram of Euclidean dynamical triangulations shown in Fig. 3 gives an example of what can happen. The picture turns out to be essentially the same in both three and four dimensions: the model possesses infinite-volume limits everywhere along the critical line $k_3^{\mathrm{crit}}(k_0)$, which fixes the bare cosmological constant as a function of the inverse Newton constant $k_0 \sim G_N^{-1}$. Along this line, there is a critical point k_0^{crit} (which we now know to be of first order in $d = 3, 4$) below which geometries generically have a very large effective or Hausdorff dimension. (In terms of geometry, this means that there are a few vertices at which the entire space-time "condenses" in the sense that almost every other vertex in the simplicial space-time is about one link-distance away from them.) Above k_0^{crit} we find the opposite phenomenon of "polymerization": a typical element contributing to the state sum is a thin branched polymer, with one or more dimensions "curled up" (an image familiar to string theorists!) such that its effective dimension is around two.

Why this happens was, at least until recently, less clear, although it has sometimes been related to the so-called conformal-factor problem. This problem has to do with the fact that the gravitational action is unbounded below, causing potential havoc in Euclidean versions of the path integral. This will be discussed in more detail below in Sect. 5.2, but it does lead directly to the next point. Namely, what all the above-mentioned approaches have in common is that they work from the outset with *Euclidean* geometries, and associated Boltzmann-type weights $\exp(-S^{\mathrm{eu}})$ in the path integral. In other words, they integrate over "space-times" which know nothing about time, light cones and causality. This is done mainly for technical reasons, since it is difficult to set up simulations with complex weights and since until recently a suitable Wick rotation was not known.

"Lorentzian dynamical triangulations", first proposed in [13] and further elaborated in [14,15] tries to establish a logical connection between the fact that non-perturbative path integrals were constructed for Euclidean instead of Lorentzian geometries and their apparent failure to lead to an interesting continuum theory. Is it conceivable that we can kill two birds with one stone, ie. cure the problem of degenerate quantum geometry by taking a path integral over geometries with a physical, Lorentzian signature? Remarkably, this is indeed what happens in the quantum gravity theories in $d < 4$ which have already been studied extensively. The way in which Lorentzian dynamical triangulations overcome the problems mentioned above is the subject of the Sect. 5.

4 Geometry from Simplices

The use of simplicial methods in general relativity goes back to the pioneering work of Regge [16]. In classical applications one tries to approximate a classical space-time geometry by a triangulation, that is, a piecewise linear space obtained by gluing together flat simplicial building blocks, which in dimension d are d-dimensional generalizations of triangles. By "flat" I mean that they are isometric to a subspace of d-dimensional Euclidean or Minkowski space. We will only be interested in gluings leading to genuine manifolds, which therefore look locally like an R^d. A nice feature of such simplicial manifolds is that their geometric properties are completely described by the discrete set $\{l_i^2\}$ of the squared lengths

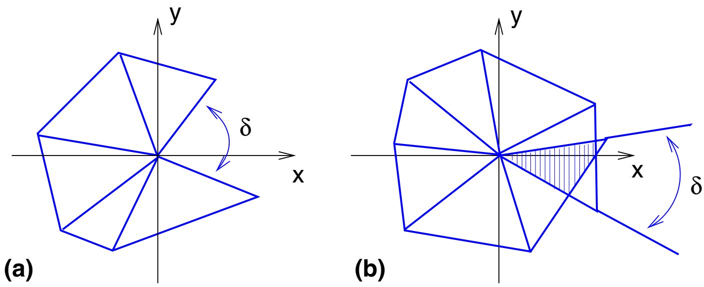


Fig. 4. Positive (**a**) and negative (**b**) space-like deficit angles δ

of their edges. Note that this amounts to a description of geometry *without the use of coordinates*. There is nothing to prevent us from re-introducing coordinate patches covering the piecewise linear manifold, for example, on each individual simplex, with suitable transition functions between patches. In such a coordinate system the metric tensor will then assume a definite form. However, for the purposes of formulating the path integral we will not be interested in doing this, but rather work with the edge lengths, which constitute a direct, regularized parametrization of the space Geom(M) of geometries.

How precisely is the intrinsic geometry of a simplicial space, most importantly, its curvature, encoded in its edge lengths? A useful example to keep in mind is the case of dimension two, which can easily be visualized. A 2d piecewise linear space is a triangulation, and its scalar curvature $R(x)$ coincides with the so-called Gaussian curvature. One way of measuring this curvature is by parallel-transporting a vector around closed curves in the manifold. In our piecewise-flat manifold such a vector will always return to its original orientation *unless* it has surrounded lattice vertices v at which the surrounding angles did not add up to 2π, but $\sum_{i\supset v}\alpha_i = 2\pi - \delta$, for $\delta \neq 0$, see Fig. 4. The so-called deficit angle δ is precisely the rotation angle picked up by the vector and is a direct measure for the scalar curvature at the vertex. The operational description to obtain the scalar curvature in higher dimensions is very similar, one basically has to sum in each point over the Gaussian curvatures of all two-dimensional submanifolds. This explains why in Regge calculus the curvature part of the Einstein action is given by a sum over building blocks of dimension $(d-2)$ which are simply the objects dual to those local two-dimensional submanifolds. More precisely, the continuum curvature and volume terms of the action become

$$\frac{1}{2}\int_{\mathcal{R}} d^d x\ \sqrt{|\det g|}^{(d)} R \longrightarrow \sum_{i\in\mathcal{R}} Vol(i^{th}\ (d-2)\text{−simplex})\ \delta_i\ , \tag{6}$$

$$\int_{\mathcal{R}} d^d x\ \sqrt{|\det g|} \qquad \longrightarrow \sum_{i\in\mathcal{R}} Vol(i^{th}\ d\text{−simplex}) \tag{7}$$

in the simplicial discretization. It is then a simple exercise in trigonometry to express the volumes and angles appearing in these formulas as functions of the edge lengths l_i, both in the Euclidean and the Minkowskian case.

The approach of dynamical triangulations uses a certain class of such simplicial space-times as an explicit, regularized realization of the space Geom(M). For a given volume N_d, this class consists of all gluings of manifold-type of a set of N_d simplicial building blocks of top-dimension d whose edge lengths are restricted to take either one or one out of two values. In the Euclidean case we set $l_i^2 = a^2$ for all i, and in the Lorentzian case we allow for both space- and time-like links with $l_i^2 \in \{-a^2, a^2\}$, where the geodesic distance a serves as a short-distance cutoff, which will be taken to zero later. Coming from the classical theory this may seem a grave restriction at first, but this is indeed not the case. Firstly, keep in mind that for the purposes of the quantum theory we want to sample the space of geometries "ergodically" at a coarse-grained scale of order a. This should be contrasted with the classical theory where the objective is usually to approximate a given, *fixed* space-time to within a length scale a. In the latter case one typically requires a much finer topology on the space of metrics or geometries. It is also straightforward to see that no local curvature degrees of freedom are suppressed by fixing the edge lengths; deficit angles in all directions are still present, although they take on only a discretized set of values. In this sense, in dynamical triangulations all geometry is in the gluing of the fundamental building blocks. This is dual to how quantum Regge calculus is set up, where one usually fixes a triangulation T and then "scans" the space of geometries by letting the l_i's run continuously over all values compatible with the triangular inequalities.

In a nutshell, Lorentzian dynamical triangulations give a definite meaning to the "integral over geometries", namely, as a sum over inequivalent Lorentzian gluings T over any number N_d of d-simplices,

$$\int_{\mathrm{Geom(M)}} \mathcal{D}[g_{\mu\nu}]\ \mathrm{e}^{iS[g_{\mu\nu}]} \quad \stackrel{\mathrm{LDT}}{\longrightarrow} \quad \sum_{T\in\mathcal{T}} \frac{1}{C_T}\ \mathrm{e}^{iS^{\mathrm{Regge}}(T)}, \tag{8}$$

where the symmetry factor $C_T = |Aut(T)|$ on the right-hand side is the order of the automorphism group of the triangulation, consisting of all maps of T onto itself which preserve the connectivity of the simplicial lattice. I will specify below what precise class $\mathcal{T}$ of triangulations should appear in the summation.

It follows from the above that in this formulation all curvatures and volumes contributing to the simplicial Regge action come in discrete units. This is again easily illustrated by the case of a two-dimensional triangulation with Euclidean signature, which according to the prescription of dynamical triangulations consists of *equilateral* triangles with squared edge lengths $+a^2$. All interior angles of such a triangle are equal to $\pi/3$, which implies that the deficit angle at any vertex v can take the values $2\pi - k_v\pi/3$, where k_v is the number of triangles meeting at v. As a consequence, the Einstein-Regge action assumes the simple form[3]

[3] Strictly speaking, the expression (9) in $d \geq 3$ is only correct for the Euclidean or the Wick-rotated Lorentzian action. In the Lorentzian case one has several types of simplices of a given dimension d, depending on how many of its links are time-like.

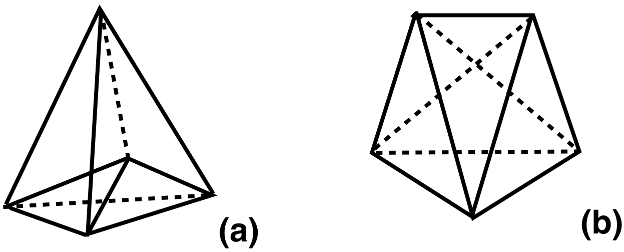


Fig. 5. The two types of Minkowskian four-simplices in four dimensions

$$S^{\mathrm{Regge}}(T) = \kappa_{d-2} N_{d-2} - \kappa_d N_d, \tag{9}$$

where the coupling constants $\kappa_i = \kappa_i(\lambda, G_N)$ are simple functions of the bare cosmological and Newton constants in d dimensions. Substituting this into the path sum in (8) leads to

$$Z(\kappa_{d-2}, \kappa_d) = \sum_{N_d} \mathrm{e}^{-i\kappa_d N_d} \sum_{N_{d-2}} \mathrm{e}^{i\kappa_{d-2} N_{d-2}} \sum_{T|_{N_d, N_{d-2}}} \frac{1}{C_T} \, . \tag{10}$$

The point of taking separate sums over the numbers of d- and $(d-2)$-simplices in (10) is to make explicit that "doing the sum" is tantamount to the combinatorial problem of *counting* triangulations of a given volume and number of simplices of co-dimension two (corresponding to the last summation in (10)).[4] It turns out that at least in two space-time dimensions the counting of geometries can be done completely explicitly, turning both Lorentzian and Euclidean quantum gravity into exactly soluble statistical models.

5 Lorentzian Nature of the Path Integral

It is now time to explain what makes our approach *Lorentzian* and why it therefore differs from previous attempts at constructing non-perturbative gravitational path integrals. The simplicial building blocks of the models are taken to be pieces of Minkowski space, and their edges have squared lengths $+a^2$ or $-a^2$. For example, the two types of four-simplices that are used in Lorentzian dynamical triangulations in dimension four are shown in Fig. 5. The first of them has four time-like and six space-like links (and therefore contains 4 time-like and 1 space-like tetrahedron), whereas the second one has six time-like and four space-like links (and contains 5 time-like tetrahedra). Since both are subspaces of flat space with signature $(-+++)$, they possess well-defined light-cone structures everywhere.

Only after the Wick rotation will all links be space-like and of equal length (see later). Nevertheless, I will use this more compact form for ease of notation.

[4] The symmetry factor C_T is almost always equal to 1 for large triangulations.

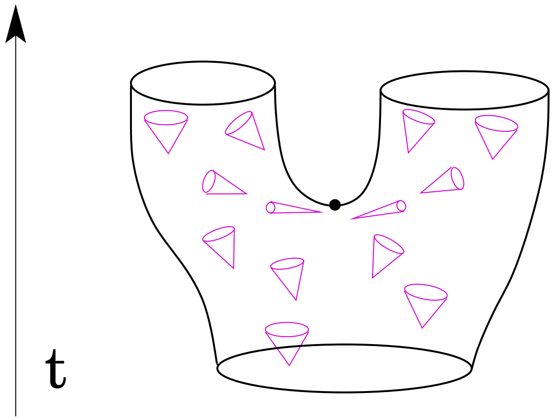


Fig. 6. At a branching point associated with a spatial topology change, light-cones get "squeezed"

In general, gluings between pairs of d-simplices are only possible when the metric properties of their $(d-1)$-faces match. Having local light cones implies causal relations between pairs of points in local neighbourhoods. Creating closed time-like curves will be avoided by requiring that all space-times contributing to the path sum possess a global "time" function t. In terms of the triangulation this means that the d-simplices are arranged such that their space-like links all lie in slices of constant integer t, and their time-like links interpolate between adjacent spatial slices t and $t+1$. Moreover, with respect to this time, we will not allow for any *spatial* topology changes[5].

This latter condition is always satisfied in classical applications, where "trouser points" like the one depicted in Fig. 6 are ruled out by the requirement of having a non-degenerate Lorentzian metric defined everywhere on M (it is geometrically obvious that the light cone and hence $g_{\mu\nu}$ must degenerate in at least one point along the "crotch"). Another way of thinking about such configurations (and their time-reversed counterparts) is that the causal past (future) of an observer changes discontinuously as her worldline passes near the singular point (see [17] and references therein for related discussions about the issue of topology change in quantum gravity).

Of course, there is no *a priori* reason in the quantum theory to not relax some of these classical causality constraints. After all, as I stressed right at the outset, path integral histories are not in general classical solutions, nor can we attribute any other direct physical meaning to them individually. It might well be that one can construct models whose path integral configurations violate causality in this strict sense, but where this notion is somehow recovered in the resulting continuum theory. What the approach of Lorentzian dynamical triangulations

[5] Note that if we were in the continuum and had introduced coordinates on space-time, such a statement would actually be diffeomorphism-invariant.

has demonstrated is that *imposing causality constraints will in general lead to a different continuum theory*. This is in contrast with the intuition one may have that "including a few isolated singular points will not make any difference". On the contrary, tampering with causality in this way is not innocent at all, as was already anticipated by Teitelboim many years ago [18].

I want to point out that one cannot conclude from the above that spatial topology changes or even fluctuations in the *space-time* topology cannot be treated in the formulation of dynamical triangulations. However, if one insists on including geometries of variable topology in a Lorentzian discrete context, one has to come up with a prescription of how to weigh these singular points in the path integral, both before and after the Wick rotation. Maybe this can be done along the lines suggested in [19]; this is clearly an interesting issue for further research.

Having said this, we next have to address the question of the Wick rotation, in other words, of how to get rid of the factor of i in the exponent of (10). Without it, this expression is an infinite sum (since the volume can become arbitrarily large) of complex terms whose convergence properties will be very difficult to establish. In this situation, a Wick rotation is simply a technical tool which – in the best of all worlds – enables us to perform the state sum and determine its continuum limit. Of course, the end result will have to be Wick-rotated back to Lorentzian signature.

Fortunately, Lorentzian dynamical triangulations come with a natural notion of Wick rotation, and the strategy I just outlined can be carried out explicitly in two space-time dimensions, leading to a unitary theory (see Sect. 5.1 below). In higher dimensions we do not yet have sufficient analytical control of the continuum theories to make specific statements about the *inverse* Wick rotation. Since we use the Wick rotation at an intermediate step, one can ask whether other Wick rotations would lead to the same result. Currently this is a somewhat academic question, since it is in practice difficult to find such alternatives. In fact, it is quite miraculous we have found a single prescription for Wick-rotating in our regularized setting, and it does not seem to have a direct continuum analogue (for more comments on this issue, see [20,21]).

Our Wick rotation W in any dimension is an injective map from Lorentzian- to Euclidean-signature simplicial space-times. Using the notation $\mathtt{T}$ for a simplicial manifold together with length assignments l_s^2 and l_t^2 to its space- and time-like links, it is defined by

$$\mathtt{T}^{\mathrm{lor}} = (T, \{l_s^2 = a^2, l_t^2 = -a^2\}) \quad \overset{W}{\mapsto} \quad \mathtt{T}^{\mathrm{eu}} = (T, \{l_s^2 = a^2, l_t^2 = a^2\}). \tag{11}$$

Note that we have not touched the connectivity of the simplicial manifold T, but only its metric properties, by mapping all time-like links of T into space-like ones, resulting in a Euclidean "space-time" of equilateral building blocks. It can be shown [15] that at the level of the corresponding weight factors in the path integral this Wick rotation[6] has precisely the desired effect of rotating to the exponentiated Regge action of the Euclideanized geometry,

$$\mathrm{e}^{iS(\mathcal{T}^{\mathrm{lor}})} \overset{W}{\mapsto} \mathrm{e}^{-S(\mathcal{T}^{\mathrm{eu}})}. \tag{12}$$

The Euclideanized path sum after the Wick rotation has the form

$$\begin{aligned} Z^{\mathrm{eu}}(\kappa_{d-2},\kappa_d) &= \sum_T \frac{1}{C_T}\,\mathrm{e}^{-\kappa_d N_d(T)+\kappa_{d-2}N_{d-2}(T)} \\ &= \sum_{N_d} \mathrm{e}^{-\kappa_d N_d} \sum_{T|_{N_d}} \frac{1}{C_T}\,\mathrm{e}^{\kappa_{d-2}N_{d-2}(T)} \\ &= \sum_{N_d} \mathrm{e}^{-\kappa_d N_d}\,\mathrm{e}^{\kappa_d^{\mathrm{crit}}(\kappa_{d-2})N_d} \times \mathrm{subleading}(N_d). \end{aligned} \tag{13}$$

In the last equality I have used that the number of Lorentzian triangulations of discrete volume N_d to leading order scales exponentially with N_d for large volumes. This can be shown explicitly in space-time dimension 2 and 3. For $d = 4$, there is strong (numerical) evidence for such an exponential bound for *Euclidean* triangulations, from which the desired result for the Lorentzian case follows (since W maps to a strict subset of all Euclidean simplicial manifolds).

From the functional form of the last line of (13) one can immediately read off some qualitative features of the phase diagram, an example of which appeared already earlier in Fig. 3. Namely, the sum over geometries Z^{eu} converges for values $\kappa_d > \kappa_d^{\mathrm{crit}}$ of the bare cosmological constant, and diverges (ie. is not defined) below this critical line. Generically, for all models of dynamical triangulations the infinite-volume limit is attained by approaching the critical line $\kappa_d^{\mathrm{crit}}(\kappa_{d-2})$ from above, ie. from inside the region of convergence of Z^{eu}. In the process of taking $N_d \to \infty$ and the cutoff $a \to 0$, one obtains a renormalized cosmological constant Λ through

$$(\kappa_d - \kappa_d^{\mathrm{crit}}) = a^\mu \Lambda + O(a^{\mu+1}). \tag{14}$$

If the scaling is canonical (which means that the dimensionality of the renormalized coupling constant is the one expected from the classical theory), the exponent is given by $\mu = d$. Note that this construction requires a positive *bare* cosmological constant in order to make the state sum converge. Moreover, by virtue of relation (14) also the *renormalized* cosmological constant must be positive. Other than that, its numerical value is not determined by this argument, but by comparing observables of the theory which depend on Λ with actual physical measurements.[7] Another interesting observation is that the inclusion of a sum

[6] To obtain a genuine Wick rotation and not just a discrete map, one introduces a complex parameter α in $l_t^2 = -\alpha a^2$. The proper prescription leading to (12) is then an analytic continuation of α from 1 to -1 through the lower-half complex plane.

[7] The non-negativity of the renormalized cosmological coupling may be taken as a first "prediction" of our construction, which in the physical case of four dimensions is indeed in agreement with current observations.

over topologies in the discretized sum (13) would lead to a super-exponential growth of at least $\propto N_d!$ of the number of triangulations with the volume N_d. Such a divergence of the path integral cannot be compensated by an additive renormalization of the cosmological constant of the kind outlined above.

There are of course ways in which one can sum divergent series of this type, for example, by performing a Borel sum. The problem with these stems from the fact that two different functions can share the same asymptotic expansion. Therefore, the series in itself is *not* sufficient to define the underlying theory uniquely. The non-uniqueness arises because of non-perturbative contributions to the path integral which are not represented in the perturbative expansion.[8] In order to fix these uniquely, an independent, non-perturbative definition of the theory is necessary. Unfortunately, for dynamically triangulated models of quantum gravity, no such definitions have been found so far. In the context of two-dimensional (Euclidean) quantum gravity this difficulty is known as the "absence of a physically motivated double-scaling limit" [22]. The same issue has recently been revived in $d = 3$ [23], where the situation is not any better.

Lastly, obtaining an interesting continuum limit may or may not require an additional fine-tuning of the inverse gravitational coupling κ_{d-2}, depending on the dimension d. In four dimensions, one would expect to find a second-order transition along the critical line, corresponding to local gravitonic excitations. The situation in $d = 3$ is less clear, but results obtained so far indicate that no fine-tuning of Newton's constant is necessary [24,25].

Before delving into the details, let me summarize briefly the results that have been obtained so far in the approach of Lorentzian dynamical triangulations. At the regularized level, that is, in the presence of a finite cutoff a for the edge lengths and an infrared cutoff for large space-time volume, they are well-defined statistical models of Lorentzian random geometries in $d = 2, 3, 4$. In particular, they obey a suitable notion of reflection-positivity and possess selfadjoint Hamiltonians.

The crucial questions are then to what extent the underlying combinatorial problems of counting all d-dimensional geometries with certain causal properties can be solved, whether continuum theories with non-trivial dynamics exist and how their bare coupling constants get renormalized in the process. What we know about Lorentzian dynamical triangulations so far is that they lead to continuum theories of quantum gravity in dimension 2 and 3. In $d = 2$, there is a complete analytic solution, which is distinct from the continuum theory produced by Euclidean dynamical triangulations. Also the matter-coupled model has been studied. In $d = 3$, there are numerical and partial analytical results which show that both a continuum theory exists and that it again differs from its Euclidean counterpart. Work on a more complete analytic solution which would give details about the geometric properties of the quantum theory is under way. In $d = 4$, the first numerical simulations are currently being set up. The challenge here is to do this for sufficiently large lattices, to be able to perform meaningful

[8] A field-theoretic example would be instantons and renormalons in QCD.

measurements. So far, we cannot make any statements about the existence and properties of a continuum theory in this physically most interesting case.

5.1 In Two Dimensions

The two-dimensional case serves as a nice illustration of the objectives of the approach, many of which can be carried out in a completely explicit manner [13]. There is just one type of building block, a flat Minkowskian triangle with two time-like edges of squared edge lengths $l_t^2 = -a^2$ and one space-like edge with $l_s^2 = a^2$. We build up a causal space-time from strips of unit height $\Delta t = 1$ (see Fig. 7), where t is an integer-valued discrete parameter that labels subsequent spatial slices, i.e. simplicial submanifolds of codimension 1 which are constructed from space-like links only. In the two-dimensional case these subspaces are one-dimensional. We choose periodic boundary conditions, such that the spatial "universes" are topologically spheres S^1 (other boundary conditions are also possible, leading to a slight modification of the effective quantum Hamiltonian [26,27]). A spatial geometry at given t is completely characterized by its length $l(t) \in \{1, 2, 3, \dots\}$, which (in units of the lattice spacing a) is simply the number of spatial edges it contains.

One simplification occurring in two dimensions is that the curvature term in the Einstein action is a topological invariant (and that therefore does not depend on the metric), given by

$$\int_M d^2x \, \sqrt{|\det g|} R = 2\pi\chi, \tag{15}$$

where χ denotes the Euler characteristic of the two-dimensional space-time M. Since we are keeping the space-time topology fixed, the exponential of i times this term is a constant overall factor that can be pulled out of the path integral and does not contribute to the dynamics. Dropping this term, we can write the discrete path integral over 2d simplicial causal space-times as

$$G_\lambda(l_{\rm in}, l_{\rm out}; t) = \sum_{\substack{\text{causal } T \\ l_{\rm in}, l_{\rm out}, t}} {\rm e}^{-i\lambda N_2} \quad \overset{\text{Wick}}{\longrightarrow} \quad \sum_{\substack{W(T) \\ l_{\rm in}, l_{\rm out}, t}} {\rm e}^{-\tilde{\lambda} N_2}, \tag{16}$$

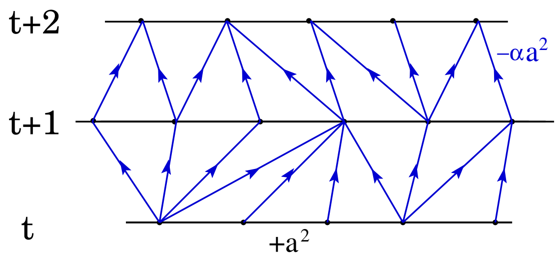


Fig. 7. Two strips of a 2d Lorentzian triangulation, with spatial slices of constant t and interpolating future-oriented time-like links

where the weight factors depend now only on the cosmological (volume) term, and $\tilde{\lambda}$ differs from λ by a finite positive numerical factor. Each history entering in the discrete propagator (16) has an in-geometry of length $l_{\rm in}$, an out-geometry of length $l_{\rm out}$, and consists of t steps. An important special case is the propagator for a single step, which in its Wick-rotated form reads[9]

$$G_{\tilde{\lambda}}(l_1,l_2;t=1)=\langle l_2|\hat{T}|l_1\rangle=\mathrm{e}^{-\tilde{\lambda}(l_1+l_2)}\sum_{T:l_1\to l_2}1\equiv\mathrm{e}^{-\tilde{\lambda}(l_1+l_2)}\frac{1}{l_1+l_2}\begin{pmatrix}l_1+l_2\\ l_1\end{pmatrix}. \tag{17}$$

The second equation in (17) defines the transfer matrix $\hat{T}$ via its matrix elements in the basis of the (improper) length eigenvectors $|l\rangle$. Knowing the eigenvalues of the transfer matrix is tantamount to a solution of the general problem by virtue of the relation

$$G_{\tilde{\lambda}}(l_1,l_2;t)=\langle l_2|\hat{T}^t|l_1\rangle. \tag{18}$$

Importantly, the propagator satisfies the composition property

$$G_{\tilde{\lambda}}(l_1,l_2;t_1+t_2)=\sum_{l=1}^{\infty}G_{\tilde{\lambda}}(l_1,l;t_1)\,l\,G_{\tilde{\lambda}}(l,l_2;t_2), \tag{19}$$

where the sum on the right-hand side is over a complete set of intermediate length eigenstates.

Next, we look for critical behaviour of the propagator $G_{\tilde{\lambda}}$ (that is, a non-analytic behaviour as a function of the renormalized coupling constant) in the limit as $a\to 0$. Since there is only one coupling, the phase diagram of the theory is just one-dimensional, and illustrated in Fig. 8. As can be read off from the explicit form of the propagator,

$$G_{\tilde{\lambda}}=\sum_{N_2}\mathrm{e}^{-\tilde{\lambda}N_2}\sum_{T|_{N_2}}1=\sum_{N_2}\mathrm{e}^{-(\tilde{\lambda}-\tilde{\lambda}^{\rm crit})N_2}\times\text{subleading}(N_2), \tag{20}$$

the discrete sum over 2d geometries converges above some critical value $\tilde{\lambda}^{\rm crit}>0$, and diverges for $\tilde{\lambda}$ below this point. In order to attain a macroscopic physical volume $\langle V\rangle:=\langle a^2N_2\rangle$ in the $a\to 0$ limit, one needs to approach $\tilde{\lambda}^{\rm crit}$ from above.

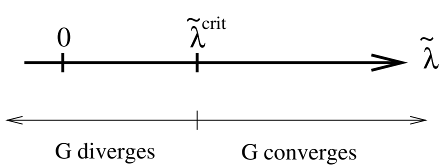


Fig. 8. The 1d phase diagram of 2d Lorentzian dynamical triangulations

[9] This is the “unmarked” propagator, see [13,11] for details.

It turns out that to get a non-trivial continuum limit, the bare cosmological coupling constant has to be fine-tuned canonically according to

$$\tilde{\lambda} - \tilde{\lambda}^{\text{crit}} = a^2 \Lambda^{\text{ren}} + O(a^3). \tag{21}$$

Note that the numerical value of $\tilde{\lambda}^{\text{crit}}$ will depend on the details of the discretization (for example, the building blocks chosen; see [26] for alternative choices), the so-called non-universal properties of the model which do not affect the quantum dynamics of the final continuum theory. At the same time, the counting variables l and t are taken to infinity while keeping the dimensionful quantities $L := al$ and $T := at$ constant. The renormalized propagator is then defined as a function of all the renormalized variables,

$$G_\Lambda(L_1, L_2; T) := \lim_{a \to 0} a^\nu G_{\tilde{\lambda}^{\text{crit}} + a^2 \Lambda}\Big(\frac{L_1}{a}, \frac{L_2}{a}; \frac{T}{a}\Big), \tag{22}$$

which also contains a multiplicative wave function renormalization. The final result for the continuum path integral of two-dimensional Lorentzian quantum gravity is obtained by an inverse Wick rotation of the continuum proper time T to iT from the Euclidean expression and is given by

$$G_\Lambda(L_{\text{in}}, L_{\text{out}}; T) = \mathrm{e}^{-\coth(i\sqrt{\Lambda}T)\sqrt{\Lambda}(L_{\text{in}} + L_{\text{out}})} \, \frac{\sqrt{\Lambda L_{\text{in}} L_{\text{out}}}}{\sinh(i\sqrt{\Lambda}T)} \, I_1\left(\frac{2\sqrt{\Lambda L_{\text{in}} L_{\text{out}}}}{\sinh(i\sqrt{\Lambda}T)}\right), \tag{23}$$

where I_1 denotes the Bessel function of the first kind.

What is the physics behind this functional expression? In two dimensions, there is not much "physics" in the sense that the *classical* Einstein equations are empty. This renders meaningless the question of a classical limit of the 2d quantum theory; whatever dynamics there is will be purely "quantum". Figure 9 shows a typical two-dimensional quantum universe: the compactified direction is "space", and the vertical axis is "time". It illustrates the typical development of the ground state of the system over time, as generated by a Monte-Carlo simulation of almost 19.000 triangles.

Since the theory has been solved analytically, we also know the explicit form of the effective quantum Hamiltonian, namely,

$$\hat{H} = -L\frac{d^2}{dL^2} - 2\frac{d}{dL} + \Lambda L. \tag{24}$$

This operator is selfadjoint on the Hilbert space $L^2(\mathbf{R}_+, LdL)$ and generates a unitary evolution in the continuum proper time T. The Hamiltonian consists of a kinetic term in the single geometric variable L (the size of the spatial universe) and a potential term depending on the renormalized cosmological constant. Its spectrum is discrete,

$$E_n = 2(n+1)\sqrt{\Lambda}, \quad n = 0, 1, 2, \ldots \tag{25}$$

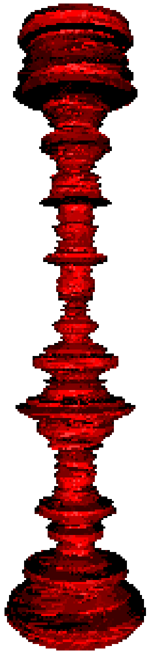


Fig. 9. A typical two-dimensional Lorentzian space-time, with volume $N_2 = 18816$ and a total proper time of $t = 168$ steps

and one can compute various expectation values, for example,

$$\langle L \rangle_n = \frac{n+1}{\sqrt{\Lambda}}, \qquad \langle L^2 \rangle_n = \frac{3}{2}\frac{(n+1)^2}{\Lambda}. \tag{26}$$

Since there is just one dimensionful constant, with $[\Lambda] = \text{length}^{-2}$, all dimensionful quantities must appear in appropriate units of Λ.

Another useful way of characterizing the continuum theory is via certain critical exponents, which in the case of gravitational theories are of a geometrical nature. The *Hausdorff dimension* d_H describes the scaling of the volume of a geodesic ball of radius R as a function of R. This very general notion can be applied to a fixed metric space, but for our purposes we are interested in the ensemble average over the entire "sum over geometries", that is, the leading-order scaling behaviour of the expectation value[10]

$$\langle V(R) \rangle \propto R^{d_H}. \tag{27}$$

The Hausdorff dimension is a truly dynamical quantity, and is *not* a priori the same as the dimensionality of the building blocks that were used to construct

[10] For the Lorentzian theory, "geodesic distance" refers to the length measurements after the Wick rotation.

the individual discrete space-times in the first place. It may even depend on the length scale of the radial distance R. Remarkably, d_H can be calculated analytically in both Lorentzian and Euclidean 2d quantum gravity (see, for example, [28]). The latter, also known as "Liouville gravity", can be obtained by performing a sum over *arbitrary* triangulated Euclidean two-geometries (with fixed topology S^2), and not just those which correspond to a Wick-rotated causal Lorentzian space-time. One finds

$$d_H = 2 \quad \text{(Lorentzian)} \qquad \text{and} \qquad d_H = 4 \quad \text{(Euclidean)}. \tag{28}$$

The geometric picture associated with the non-canonical value of d_H in the Euclidean case is that of a fractal geometry, with wildly branching "baby universes". This branching behaviour is incompatible with the causal structure required in the Lorentzian case, and the geometry of the Lorentzian quantum ground state is much better behaved, although it is by no means smooth as we have already seen.

We conclude that the continuum theories of 2d quantum gravity with Euclidean and Lorentzian signature are distinct. They can be related by a somewhat complicated renormalization procedure which one may think of as "integrating out the baby universes" [29], which is not at all as simple as "sticking a factor of i in the right place". In a way, this is not unexpected in view of the fact that (the spaces of) Euclidean and Lorentzian geometries are already classically very different objects. I am not claiming that from the point of view of 2d quantum gravity, one signature is better than the other. This seems a matter of taste, since neither theory describes any aspects of real nature. Nevertheless, what we have shown is that imposing causality constraints at the level of the individual histories in the path integral changes the outcome radically, a feature one may expect to generalize to higher dimensions.

Let me comment at this point about the role of the integer t which labels the time steps in the propagator (18) and its higher-dimensional analogues. In the first place, it is one of the many discrete parameters that label the regularized space-times *in a coordinate-invariant way.* In any given Minkowskian building block, one may introduce proper-time coordinates whose value coincides (up to a constant factor depending on the type of the building block) with the discrete time t on the spatial slices. However, this is where the analogue with continuum proper time ends, since it is in general impossible to extend such coordinate patches over more than one time step, because of the presence of curvature singularities. Next, there is no claim that the propagator with respect to t or its continuum analogue T has a distinguished *physical* meaning, despite being invariantly defined. Nevertheless, we do believe strongly that it contains all physical information about the "quantum geometry". In other words, all observables and propagators (which may depend on other notions of "time") can in principle be computed from our propagator in t.[11] This can of course be difficult in practice, but this is only to be expected.

[11] A related result has already been demonstrated for the proper-time propagator in two-dimensional *Euclidean* quantum gravity [30].

Coming from Euclidean quantum gravity, there are specific reasons for looking at the behaviour of the matter-coupled theory in two dimensions. The coupling of matter fields to Lorentzian dynamical triangulations can be achieved in the usual manner by including for each given geometry T in the path integral a summation over all matter degrees of freedom on T, resulting in a double sum over geometric and matter variables. For example, adding Ising spins to 2d Lorentzian gravity is described by the partition function

$$Z(\lambda, \beta_I) = \sum_{N_2} \mathrm{e}^{-\lambda N_2} \sum_{\substack{\text{causal} \\ T \in \mathcal{T}_{N_2}}} \sum_{\{\sigma_i = \pm 1\}} \mathrm{e}^{\frac{\beta_I}{2} \sum_{<ij>} \sigma_i \sigma_j}, \tag{29}$$

where the last sum on the right is over the spin configurations of the Ising model on the triangulation T. The analogous model on Euclidean triangulations has been solved exactly [31], and its continuum matter behaviour is characterized by the critical exponents

$$\alpha = -1, \quad \beta = 0.5, \quad \gamma = 2, \qquad \text{(Euclidean)} \tag{30}$$

for the specific heat, the magnetization and the magnetic susceptibility respectively. These differ from the ones found for the Ising model on a fixed, flat lattice, the so-called Onsager exponents. The transition here is third-order, reflecting the influence of the fractal background on which the matter is propagating.

The same Ising model, when coupled to Lorentzian geometries according to (29), has not so far been solved exactly, but its critical matter exponents have been determined numerically and by means of a diagrammatic high-T expansion [32] and agree (within error bars) with the Onsager exponents, that is,

$$\alpha = 0, \quad \beta = 0.125, \quad \gamma = 1.75. \qquad \text{(Lorentzian)} \tag{31}$$

So, interestingly, despite the fluctuations of the geometric ensemble evident in Fig. 9, the conformal matter behaves *as if* it lived on a static flat lattice. This indicates a certain robustness of the Onsager behaviour in the presence of such fluctuations. Does it also imply there cannot be any back-reaction of the matter on the geometry? In order to answer this question, Lorentzian quantum gravity was coupled to "a lot of matter", in this case, eight copies of Ising models [33]. The partition function is a direct generalization of (29). For a given triangulation, there are 8 independent Ising models, which interact with each other only via their common interaction with the ensemble of geometries.

Looking again at a typical "universe", depicted in Fig. 10, its geometry is now significantly changed in comparison with the case without matter. Part of it is squeezed down to a spatial universe of minimal size, with the remainder forming a genuinely extended space-time. A measurement of the critical behaviour of the matter on this piece of the universe again produces values compatible with the Onsager exponents![12] This is a very interesting result from the point of view

[12] The same would of course not hold for the degenerate part of the space-time which is effectively one-dimensional.

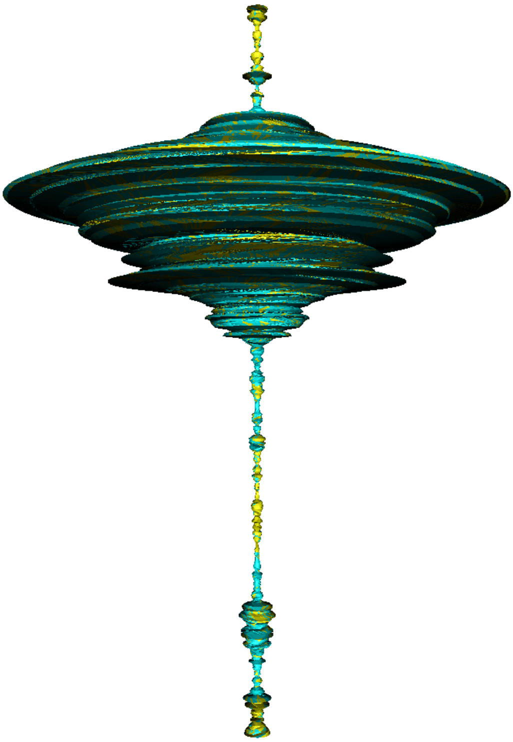


Fig. 10. A typical two-dimensional Lorentzian geometry in the presence of eight Ising models, for volume $N_2 = 73926$ and a total proper time $t = 333$

of Liouville gravity, which does not seem to produce meaningful matter-coupled models beyond a central charge of one, the famous $c = 1$ barrier. (A model with n Ising spins corresponds to central charge $c = n/2$.) We conclude that *causal* space-times are better carrier spaces for matter fields in 2d quantum gravity.

5.2 In Three Dimensions

Having discovered the many beautiful features of being Lorentzian in two dimensions, the next challenge is to solve the dynamically triangulated model in three dimensions and understand the geometric properties of the continuum theory it gives rise to. This will bring us a step closer to our ultimate goal, the four-dimensional quantum theory.

Despite its reputation as an "exactly soluble theory", many aspects of quantum gravity in 2+1 dimensions remain to be understood. There is still an unresolved tension between (i) the gauge (Chern-Simons) formulation in which the constraints can be solved in a straightforward way before or after quantization, leading to a quantized finite-dimensional phase space, and (ii) a path integral formulation in terms of "$g_{\mu\nu}$" which seems just about as intractable as the four-dimensional theory, and is power-counting non-renormalizable.

Since Lorentzian dynamical triangulations are really a regularized and non-perturbative version of the latter, a solution of the model should help to bridge this gap. Part of the trouble with gravitational path integrals is the "conformal-factor problem", which makes its first appearance in $d = 3$.[13] The conformal part of the metric, ie. the mode associated with an overall scaling of all components of the metric tensor, contributes to the action with a kinetic term of the wrong sign. This is most easily seen by considering just the curvature term of the Einstein action,

$$S = \int d^d x \sqrt{g}(R + \dots), \tag{32}$$

and performing a conformal transformation $g_{\mu\nu} \to g'_{\mu\nu} = \mathrm{e}^{\phi} g_{\mu\nu}$ on the metric. This is *not* a gauge transformation and leads to a change

$$S \to S' = \int d^d x \sqrt{g'}(-(\partial_0 \phi)^2 + \dots) \tag{33}$$

in the action, with the anticipated negative kinetic term for the conformal field ϕ. In the perturbative theory, this is not a real problem since the conformal term can be isolated explicitly and eliminated. However, the ensuing unboundedness of the action spells potential trouble for any non-perturbative geometric path integral (that is either Euclidean from the outset, or has been Euclideanized by a suitable Wick rotation), since the Euclidean weight factors $\exp(-S)$ $=\exp(\dot{\phi}^2 + \dots)$ can become arbitrarily large. We will see that this problem arises in our approach too, and how it is resolved non-perturbatively.

First to some basics of Lorentzian dynamical triangulations in three dimensions. The construction of space-time manifolds is completely analogous to the 2d case. Slices of constant integer t are now two-dimensional space-like, equilateral triangulations of a given, fixed topology ${}^{(2)}\Sigma$, and time-like edges interpolate between adjacent slices t and $t+1$. The building blocks are given by two types of tetrahedra: one of them has three space-like and three time-like edges, and shares its space-like face with a slice t =const, the other has four time-like and two space-like edges, the latter belonging to two distinct adjacent spatial slices (Fig. 11). We often denote the different tetrahedral types by the numbers of vertices (n, m) they have in common with two subsequent slices, which in three dimensions can take the values (3,1) (together with its time inverse (1,3)) and (2,2). Within a given sandwich $\Delta t = 1$, a (2,2)-tetrahedron can be glued to other (2,2)'s, as well as to (3,1)- and (1,3)-tetrahedra, but a (1,3) can never be glued directly to a (3,1), since their triangular faces do not match.

[13] A more detailed account of the history of this problem in quantum gravity can be found in [20].

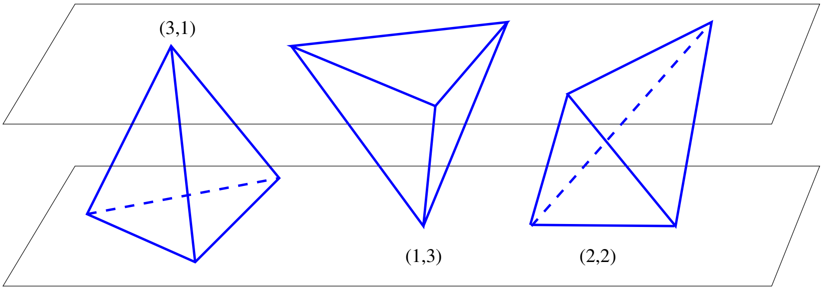


Fig. 11. The three types of tetrahedral building blocks used in 3d Lorentzian gravity

The simplicial action after the Wick rotation reads

$$S = -\kappa_1 N_1(T) + \kappa_3 N_3(T) \equiv N_3(T)\Big(-\kappa_1 \frac{N_1(T)}{N_3(T)} + \kappa_3\Big), \tag{34}$$

where the latter form is useful in the discussion of Monte-Carlo simulations, which are usually performed at (approximately) constant volume. The phase structure of the 3d model with spherical spatial topology, ${}^{(2)}\Sigma = S^2$, has been determined with the help of numerical simulations [24]. As expected, there is a critical line $\kappa_3^{\mathrm{crit}}(\kappa_1)$. After fine-tuning to this line, there is no further phase transition[14] along it as a function of the inverse Newton coupling κ_1.

Where is our conformal-mode problem? If we keep the total volume N_3 fixed, the Euclidean action is not actually unbounded, but because of the nature of our regularization restricted by the range of the "order parameter" $\xi := N_1/N_3$ which kinematically can only take values in the interval $[1, 5/4]$ [15]. This by no means implies we have removed the problem by hand. Firstly, one can explicitly identify configurations which minimize the action (34) and, secondly, the unboundedness could well be recovered upon taking the continuum limit. However, what happens dynamically is that even in the continuum limit (as far as can be deduced from the simulations [24,35]), ξ stays bounded away from its "conformal maximum", which means that the quantum theory of Lorentzian 3d gravity is *not* dominated by the dynamics of the conformal mode. Configurations with minimal action exist, but they are entropically suppressed. This is clearly a non-perturbative effect which involves not just the action, but also the "measure" of the path integral. A similar argument of a non-perturbative cancellation between certain Faddeev-Popov determinants and the conformal divergence can be made in a gauge-fixed continuum computation [20].[15]

[14] The first simulations did report a first-order transition at large κ_1, but this was presumably a numerical artefact; upon slightly generalizing the class of allowed geometries, this transition has now disappeared [34].

[15] Of course, since the continuum path integral cannot really be *done* (strictly speaking, not even in two dimensions), the cancellation argument has to rely on certain (plausible) assumptions about the behaviour of the path integral under renormalization.

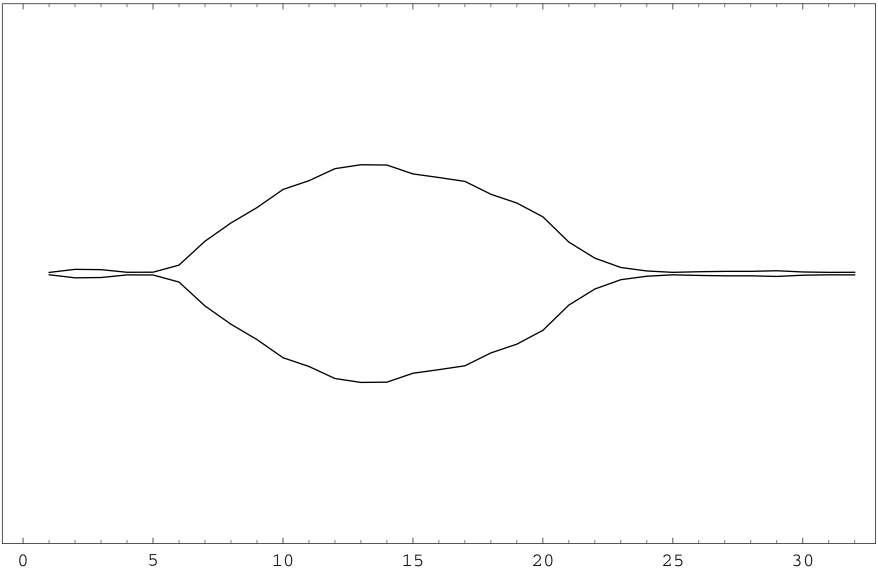


Fig. 12. A typical three-dimensional universe, represented as a distribution of two-volumes $N_2(t)$ of spatial slices at proper times $t \in [0, 32]$, at $k_0 = 5.0$

This result is reassuring, because it shows that (Euclideanized) path integrals are not doomed to fail, if only they are set up properly and non-perturbatively. It also agrees with the expectation one has from canonical treatments of the theory where it is obvious that the conformal mode is not a propagating degree of freedom.

What can we say about the quantum dynamics of 3d Lorentzian gravity and the geometry of its ground state? Fig. 12 shows a snapshot of a typical "universe" produced by the Monte-Carlo simulations. The only variable plotted as a function of the discrete time t is the two-volume of a spatial slice. What has been determined are the macroscopic scaling properties of this universe; they are in agreement with those of a genuine three-dimensional compact space-time, its time extent scaling $\propto N_3^{1/3}$ and its spatial volume $\propto N_3^{2/3}$.

Current efforts are directed at trying to analyze the detailed microscopic geometric properties of the quantum universe, its effective quantum Hamiltonian, and at gaining an explicit analytic understanding of the conformal-factor cancellation. One important question is how exactly the conformal mode decouples from a propagator like $G(g^{(\text{in})}, g^{(\text{out})})$, although it appears among the labels parametrizing the in- and out-geometries g. One does not in general expect to be able to make much progress in solving a three-dimensional statistical model analytically. However, we anticipate some simplifying features in the case of pure three-dimensional gravity, which is known to describe the dynamics of a finite number of physical parameters only.

There are two main strands of investigation, one for space-times $\mathbf{R} \times S^2$ and using matrix model techniques, and the other for space-times $\mathbf{R} \times T^2$ with flat toroidal spatial slices. An observation that is being used in both is the fact that the combinatorics of the transfer matrix, crucial to the solution of the full problem, is encoded in a *two*-dimensional graph. The transfer matrix $\hat{T}$, defined in analogy with (17), describes all possible transitions from one spatial 2d triangulation to the next. Such a transition is nothing but a three-dimensional sandwich geometry $[t, t+1]$, and is completely characterized by the two-dimensional pattern that emerges when one intersects this geometry at the intermediate time $t+1/2$. One associates with each time-like triangle a coloured edge where the triangle meets the $(t+1/2)$-surface. A blue edge belongs to a triangle whose base lies in the triangulation at time t, and a red edge denotes an upside-down triangle with base at $t+1$. The intersection pattern can therefore be viewed as a combined tri- and quadrangulation, made out of red triangles, blue triangles, and squares with alternating red and blue sides.

Graphs of this type, or equivalently their duals, are also generated by the large-N limit of a hermitian two-matrix model with partition function

$$Z(\alpha_1, \alpha_2, \beta) = \int dA_{N\times N}\ dB_{N\times N}\ \mathrm{e}^{-N\ \mathrm{Tr}(\frac{1}{2}A^2+\frac{1}{2}B^2-\alpha_1 A^3-\alpha_2 B^3-\beta ABAB)}. \quad (35)$$

The cubic and quartic interaction terms in the exponent correspond to the tri- and four-valent intersections of the dual bi-coloured spherical graph characterizing a piece of space-time. In fact, as was shown in [36], the matrix model gives an embedding of the gravitational model we are after, since it generates *more* graphs than those corresponding to regular three-dimensional geometries. Interestingly, from a geometric point of view these can be interpreted as wormhole configurations. Some explicit examples are shown in Fig. 13; the graphs consist of squares since they are taken from a "pyramid" variant of three-dimensional gravity, cf. footnote 16. Blue and red edges are in these pictures represented by solid and dashed lines.

The matrix model has been solved analytically for the diagonal case $\alpha_1 = \alpha_2$ [37], and its second-order phase transition separates the phase where wormholes are rare from that where they are abundant.[16] One therefore concludes that Lorentzian gravity as given by dynamical triangulations should correspond to the former.

It turns out that to extract information about the quantum Hamiltonian of the system, one must consider the off-diagonal case where the two α-couplings are different. Only in that case can one distinguish which part of the intersection graph comes from "below" (time t) and which from "above" (time $t+1$). The colouring of the two-dimensional graph is really the memory of the original three-dimensional nature of the problem. It turns out that even for α's which

[16] More precisely, these results apply to a variant of (35) where the cubic terms A^3 and B^3 have been replaced by quartic terms A^4 and B^4. Geometrically, this corresponds to using pyramids instead of the tetrahedral building blocks, a difference that is unlikely to affect the continuum theory.

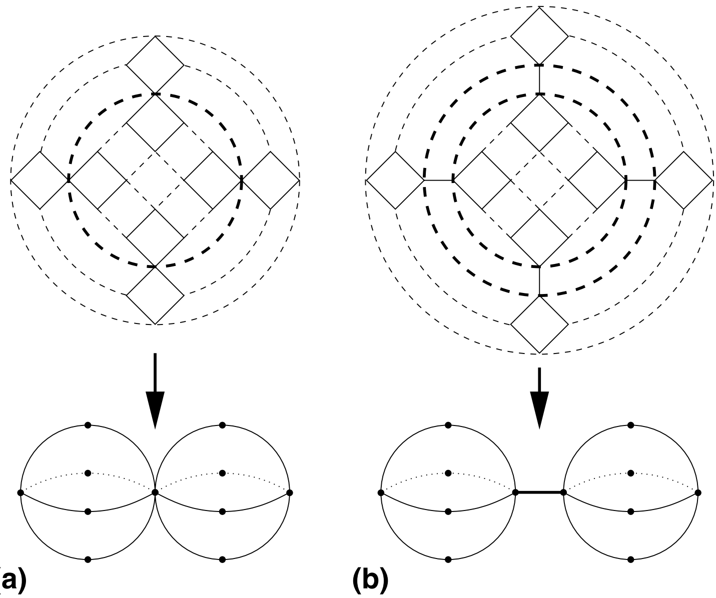


Fig. 13. Examples of quadrangulations at $t+1/2$ corresponding to wormholes at time t. Shrinking the dashed links to zero, one obtains the two-geometries at the bottom. The *thick dashed lines* at the top are contracted to points where wormholes begin or end

differ only infinitesimally, this is a highly non-trivial problem. Making a natural ansatz for the analytic structure of the eigenvalue densities that appear in the partition function, a consistent set of equations has now been found, which will hopefully yield more details about the effective Hamiltonian of the quantum system [38]. Since there are no non-trivial Teichmüller parameters in the sphere case, what one might expect on dimensional grounds is a differential operator in the two-volume V_2 of the kind [34]

$$\hat{H} = -c_1 G_N V_2 \frac{d^2}{dV_2^2} - c_2 \Lambda V_2, \tag{36}$$

where the c_i are numerical constants.

A second direction of attack are *cosmological models* of 3d gravity. They are symmetry-reduced in the sense that only a restricted class of spatial geometries is allowed at integer values of t, and also additional conditions may be imposed on the interpolating three-dimensional Lorentzian geometries. All models studied so far have flat tori as their spatial slices, the simplest case with a non-trivial physical configuration space, spanned by two real Teichmüller parameters (apart from the two-volume of the spatial slices). Flat two-dimensional tori can be obtained by suitably identifying the boundaries of a piece of the triangulated plane. Since we are working with equilateral triangles, this amounts to a piece of regular triangulation where exactly six triangles meet at every (interior) vertex point.

Even if the spatial slices have been chosen as spaces of constant curvature, this still leaves a number of possibilities of how the space-time in between can be filled in. One extreme choice would be to allow any intermediate three-geometry. By this we would probably not gain much in terms of simplifying the model, which obviously is a major motivation behind going "cosmological". By contrast, the first model studied had very simple interpolating geometries. The most transparent realization of this model is in terms of (4,1)- and (1,4)-pyramids rather than the (3,1)- and (1,3)-tetrahedra (a modification we already encoutered in the discussion of the matrix model), so that the spatial slices at integer-t are regular square lattices [39]. The corresponding 2d building blocks of the intersection graph at half-integer t are now blue squares, red squares and – as before – red-and-blue squares. If the (cut-open) tori at times $t_1 = t$ and $t_2 = t + 1$ consist of l_i columns and m_i rows, $i = 1, 2$, any allowed intersection pattern is a rectangle of size $(l_1 + l_2) \times (m_1 + m_2)$. An example is shown in Fig. 14. The trouble with this simple model is that it does not have enough entropy: the number of possible interpolating sandwiches between two neighbouring spatial slices is given by

$$\text{entropy} \propto \binom{l_1 + l_2}{l_1} \binom{m_1 + m_2}{m_1}, \tag{37}$$

which is roughly speaking the square of the entropy of the two-dimensional Lorentzian model, cf. (17). This is not enough in the sense that the number of "microstates" in a piece of space-time $\Delta t = 1$ scales asymptotically only with the linear size of the tori, ie. like exp(c·length). Such a behaviour cannot "compete" with the exponential damping exp(c'·area) coming from the cosmological term

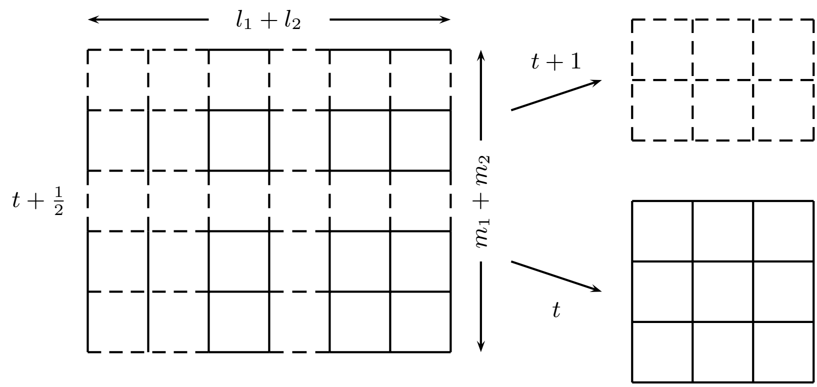


Fig. 14. The cosmological "pyramid model" has regular slices at both integer and half-integer times.

in the action. Thus, the only space-times that will not be exponentially damped in the continuum limit will be those whose spatial slices are essentially one-dimensional. This clearly is a limit that has nothing to do with the description of 3d quantum geometries we are after. In particular, the model is unsuitable for studying the conformal-mode cancellation.

I have included a discussion of this model because it suggests a potential problem for the path integral in models that impose severe symmetry constraints *before* quantization. Prime examples of this are continuum mini-superspace models with only a finite number of dynamical degrees of freedom, whose path integral formulations are riddled with difficulties. Lorentzian dynamically triangulated models are more flexible concerning the imposition of such constraints.

The next cosmological model I will consider has also flat tori at integer-t, but allows for more general geometries in between the slices. As a consequence, it does not suffer from the problem described above. The easiest way of describing the geometry of this so-called *hexagon model* is by specifying the intersection patterns at half-integer t. One such pattern can be thought of as a tiling of a regular piece of a flat equilateral triangulation with three types of coloured rhombi. The colouring of the rhombi again encodes the orientation in three dimensions of the associated tetrahedral building block. A blue rhombus stands for a pair of (3,1)-tetrahedra, glued together along a common time-like face, a red rhombus for a pair of (1,3)-tetrahedra, and the rhombus with alternating blue and red sides is a (distorted) representation of a (2,2)-tetrahedron. Opposite sides of the regular triangular "background lattice" are to be identified to create the topology of a two-torus. The beautiful feature of this model is the fact that any complete tiling of this lattice by matching rhombic tiles *automatically* gives rise to flat two-tori on the two spatial boundaries of the associated sandwich $[t, t+1]$ [40].

After the Wick rotation, the one-step propagator of this model can be written as

$$G(g^{(1)}, g^{(2)}; \varDelta t = 1) \equiv \langle g^{(2)} | \hat{T} | g^{(1)} \rangle = \mathcal{C}(g^{(1)}, g^{(2)})\ \mathrm{e}^{-S(g^{(1)}, g^{(2)})}. \qquad (38)$$

We note here a distinguishing property of the hexagon model, namely, a factorization of G into a Boltzmann weight $\exp(-S)$ and a combinatorial term $\mathcal{C}$ which counts the number of distinct sandwich geometries with fixed toroidal boundaries $g^{(1)}$ and $g^{(2)}$, both of which depend on the boundary data only, and not on the details of the three-dimensional triangulation of its interior. The leading asymptotics of the entropy term is determined by the combinatorics of a model of so-called *vicious walkers*. The walkers are usually represented by an ensemble of paths that move up a tilted square lattice, taking steps either diagonally to the left or to the right, in such a way that at most one path passes through any one lattice vertex.

The paths of the hexagon model are sequences of rhombi that have been put down on the background lattice so they lie on one of their sides (types B and C in Fig. 15). Because of the toroidal boundary conditions, such B-C-paths wind around the background lattice in the "vertical direction" (on figures such

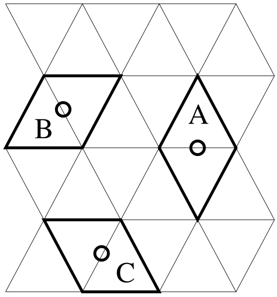


Fig. 15. A rhombus can be put onto the triangular background lattice with three different orientations, A, B or C

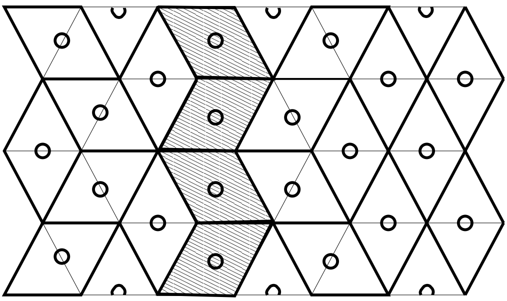


Fig. 16. An example of a periodic tiling of the triangular background lattice. The shaded region is a B-C-path with winding number (0,1)

as Fig. 16), which for the purposes of solving the 2d statistical model of vicious walkers we may think of as the time direction. The transfer matrix of this model can be diagonalized explicitly. Let us denote the number of vicious-walker paths by $w/2$, the width of the background lattice by $l+w$ and its height (in time direction) by m, all in lattice units. It turns out that for the simplest version of the model we can set $m=l$ without loss of generality. We are now interested in the number $\mathcal{N}(l,w)$ which solves the following combinatorial problem:

> Given two even integers l and w, how many ways $\mathcal{N}(l,w)$ are there of drawing $w/2$ non-intersecting paths of winding number $(0,1)$ (in the horizontal and vertical direction) onto a tilted square lattice of width $l+w$ and height l, with periodic boundary conditions in both directions?

Denoting by $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{w/2})$, $\lambda_i \in \{0, 2, 4, \dots, l+w \equiv 0\}$, the vector of positions of the vicious walkers along the horizontal axis, the eigenvectors of the transfer matrix have the form

$$\Psi(\boldsymbol{\lambda}) = \frac{1}{\sqrt{\frac{w}{2}!}} \det[z_j^{\lambda_i}], \qquad 1 \le i, j \le \frac{w}{2}, \tag{39}$$

where the complex numbers z_j are given by

$$z_j = e^{i\pi \frac{k_j}{l+w}} e^{i\pi \frac{w-2}{l+w}}, \quad 0 \le k_1 < k_2 \cdots < k_{w/2} \le \frac{l+w}{2} - 1. \tag{40}$$

This result can be understood by observing that for a single walker in the same representation, taking a step to the right (left) is represented by a multiplication (division) by z, that is,

$$\Psi(\lambda) = z^\lambda \quad \Longrightarrow \quad z\Psi(\lambda) = z^{\lambda+1} \equiv \Psi(\lambda+1). \tag{41}$$

The expression (39) is an appropriately antisymmetrized and normalized version for the case of several walkers. In this representation, the transfer matrix[17] takes the form

$$\hat{T}_{\mathrm{VW}} = \prod_{i=1}^{w/2} (\frac{1}{z_i} + 2 + z_i). \tag{42}$$

The final result in the limit as both $l, w \to \infty$, with a fixed ratio $\alpha := \frac{w}{l+w}$, is to leading order given by

$$\mathcal{N}(l,w) = C(\alpha)^{\frac{lw}{2}}, \qquad C(\alpha) = \exp\left[\frac{2}{\alpha} \int_0^{\alpha/2} dy \ \log(2\cos \pi y)\right]. \tag{43}$$

This shows that the hexagon model has indeed enough entropy, since the number of possible intermediate geometries scales exponentially with the area, and not just with the linear dimension of the tori involved.

Another attractive feature of the model is that the Teichmüller parameters $\tau(t) = \tau_1(t) + i\tau_2(t)$ of the spatial tori at time t can be written explicitly as functions of the discrete variables describing the Lorentzian simplicial spacetime. It turns out that the real parameter τ_1 is not dynamical, so that the wave functions of the model are labelled by just two numbers, the two-volume $v(t)$ and $\tau_2(t)$.[18] Expanding the euclideanized action for small $\Delta t = a$, one finds

$$S = \tilde{\lambda} v - \tilde{k} a^2 v \Big(\Big(\frac{\dot{v}}{v}\Big)^2 - \Big(\frac{\dot{\tau}_2}{\tau_2}\Big)^2 \Big) + \ldots, \tag{44}$$

where $\tilde{\lambda}$ and $\tilde{k}$ are proportional to the bare cosmological and inverse Newton's constants. This has the expected modular-invariant form, with a standard kinetic term for τ_2, and one with the wrong sign for the area v. Of course, this is our old friend, the (global) conformal mode!

What we are after is the "effective action", containing contributions from both (44) and the state counting, namely,

$$S^{\mathrm{eff}} := S - \log(\mathrm{entropy}) = v(\tilde{\lambda} - C) + ??? \tag{45}$$

[17] This is the transfer matrix corresponding to a "double step" in time; a single step would lead to a position vector with odd λ_i's.

[18] The model can be generalized to have non-trivial τ_1 by allowing for B-C-paths with higher winding numbers [41].

In order to say anything about the cancellation or otherwise of the conformal divergence, we need more than just the leading-order term (43) of the entropy of the hexagon model. Unlike the exponential term, these subleading terms are sensitive to the colouring of the intersection graph, and efforts are under way to solve the corresponding vicious-walker problem [41].

5.3 Beyond Three Dimensions

As already mentioned earlier, there is not much to report at this stage on the nature of the continuum limit in the physical case of four dimensions. The first Monte-Carlo simulations are just being set up, but any conclusive statements are likely to involve a combination of analytical and numerical arguments. Also it should be kept in mind that, unlike in previous simulations of four-dimensional *Euclidean* dynamical triangulations, the space-times involved here are not isotropic. Measurements of two-point functions, say, will be sensitive to whether the distances are time- or space-like, and therefore more computing power will be necessary to achieve a statistics comparable to the Euclidean case.

One way of making progress in four dimensions will be by studying geometries with special symmetries, along the lines of the 3d cosmological models discussed above. It should be noted that popular symmetry reductions, such as spherical or cylindrical symmetry, cannot be implemented *exactly* because of the nature of our discretization. They can at best be realized approximately, which in view of the results of the previous subsection may be a good thing since it will ensure that a sufficient number of microstates contributes to the state sum. An important application in this context is the construction of a path integral for spherical black hole configurations. Already the formulation of the problem has a number of challenging aspects, for example, the inclusion of non-trivial boundaries, an explicit realization of the (near-)spherical symmetry, and of a "horizon finde", some of which have been addressed and solved in [42,43]. It will be extremely interesting to see what Lorentzian dynamical triangulations have to say about the famous thermodynamic properties of quantum black holes from a non-perturbative point of view. These questions are currently under study.

6 Brief Conclusion

As we have seen, the method of *Lorentzian dynamical triangulations* constitutes a well-defined regularized framework for constructing non-perturbative theories of quantum gravity. Technically, they can be characterized as regularized sums over simplicial random geometries with a time arrow and certain causality properties. In dimension $d < 4$, interesting continuum limits have been shown to exist. Their geometric properties have been explored, almost exhaustively in two, and partly in three dimensions. Both are examples of Lorentzian quantum gravitational theories which as continuum theories are *inequivalent* to their Euclidean counterparts, and the relation between the two is *not* that of some simple analytic continuation of the form $t \mapsto it$. The origin of the discrepancy between quantum gravity with Euclidean and Lorentzian signature lies in the absence of causality-violating branching points for geometries in the latter. Since in dimension $d \geq 3$, the approach of *Euclidean* dynamical triangulations seems to have serious problems, I am greatly encouraged by the fact that the 3d Lorentzian model is better behaved. Of course, it still needs to be verified explicitly that the imposition of causality conditions is indeed the correct remedy to cure the *four*-dimensional theory of its apparent diseases. One step in that direction will be to show that the non-perturbative cancellation mechanism for the conformal divergence is also present in $d = 4$.

Two warnings may be in order at this point. Firstly, there is a priori nothing *discrete* about the quantum gravitational theories this method produces. Its "discreteness" refers merely to the intermediate regularization that was chosen to make the non-perturbative path sums converge.[19] In particular, there is nothing in the construction suggesting the presence of any kind of "fundamental discreteness", as has been found in canonical models of four-dimensional quantum gravity [44–46]. Secondly, one should refrain from trying to interpret the discrete expressions of the regularized model as some kind of approximation of the "real" quantum theory *before* one has shown the existence of a continuum limit which (at least in dimension four) is an interacting theory of geometric degrees of freedom.

In conclusion, I have described here a possible path for constructing a non-perturbative quantum theory of gravity, by applying standard tools from both quantum field theory and the theory of critical phenomena to theories of fluctuating geometry. Investigation of the continuum theories in two and three spacetime dimensions has already led to exciting new insights into the relation between the Lorentzian and Euclidean quantum theories, and ways of understanding and resolving the conformal sickness of gravitational path integrals, as well as bringing in new tools from combinatorics and statistical mechanics. I hope this has convinced you that the method of Lorentzian dynamical triangulations stands a good chance of throwing some light on the ever-elusive quantization of general relativity!

Acknowledgements

I would like to thank D. Giulini, C. Kiefer and C. Lämmerzahl for organizing an interesting and inspiring meeting at Bad Honnef, and for inviting me to speak there. Support through the EU network on "Discrete Random Geometry", grant HPRN-CT-1999-00161, and the ESF network no.82 on "Geometry and Disorder" is also gratefully acknowledged.

---

[19] It should maybe be emphasized that there are precious few methods around that get even that far, and possess a coordinate-invariant cutoff.

References

1. J. Ambjørn: *Simplicial Euclidean and Lorentzian quantum gravity*, plenary talk given at GR16 [gr-qc/0201028]
2. R. Loll: *Discrete Lorentzian quantum gravity*, Nucl. Phys. B (Proc. Suppl.) 94 (2001) 96–107 [hep-th/0011194]
3. C. Rovelli: *Loop quantum gravity*, Living Rev. Rel. 1 (1998), http://www.livingreviews.org [gr-qc/9710008]
4. T. Thiemann: *Introduction to modern canonical quantum general relativity* [gr-qc/0110034]
5. D. Oriti: *Spacetime geometry from algebra: Spin foam models for non-perturbative quantum gravity*, Rept. Prog. Phys. 64 (2001) 1489–1544 [gr-qc/0106091]
6. C. Grosche and F. Steiner, *Handbook of Feynman path integrals*, Springer tracts in modern physics 145, Springer, Berlin, 1998
7. M. Reed and B. Simon, *Methods of modern mathematical physics, vol.2: Fourier analysis, self-adjointness*, Academic Press, San Diego, 1975
8. S.W. Hawking: in *General relativity: an Einstein centenary survey*, ed. S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979) 746–789
9. R.M. Williams: *Recent progress in Regge calculus*, Nucl. Phys. B (Proc. Suppl.) 57 (1997) 73–81 [gr-qc/9702006]
10. J. Ambjørn, B. Durhuus and T. Jonsson, *Quantum geometry*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, UK, 1997
11. J. Ambjørn, J. Jurkiewicz and R. Loll: Lorentzian and Euclidean quantum gravity – analytical and numerical results, in: *M-Theory and Quantum Geometry*, eds. L. Thorlacius and T. Jonsson, NATO Science Series (Kluwer Academic Publishers, 2000) 382–449 [hep-th/0001124]
12. R. Loll: *Discrete approaches to quantum gravity in four dimensions*, Living Reviews in Relativity 13 (1998), http://www.livingreviews.org [gr-qc/9805049]
13. J. Ambjørn and R. Loll: *Non-perturbative Lorentzian quantum gravity, causality and topology change*, Nucl. Phys. B 536 (1998) 407–434 [hep-th/9805108]
14. J. Ambjørn, J. Jurkiewicz and R. Loll: *A nonperturbative Lorentzian path integral for gravity*, Phys. Rev. Lett. 85 (2000) 924–927 [hep-th/0002050]
15. J. Ambjørn, J. Jurkiewicz and R. Loll: *Dynamically triangulating Lorentzian quantum gravity*, Nucl. Phys. B 610 (2001) 347–382 [hep-th/0105267]
16. T. Regge: *General relativity without coordinates*, Nuovo Cim. A 19 (1961) 558–571
17. F. Dowker: *Topology change in quantum gravity*, Contribution to the proceedings of Stephen Hawking's 60th birthday conference, Cambridge Jan 2002 [gr-qc/0206020]
18. C. Teitelboim: *Causality versus gauge invariance in quantum gravity and supergravity*, Phys. Rev. Lett. 50 (1983) 705–708
19. J. Louko and R.D. Sorkin: *Complex actions in two-dimensional topology change*, Class. Quant. Grav. 14 (1997) 179–204 [gr-qc/9511023]
20. A. Dasgupta and R. Loll: *A proper-time cure for the conformal sickness in quantum gravity*, Nucl. Phys. B 606 (2001) 357–379 [hep-th/0103186]
21. A. Dasgupta: *The real Wick rotations in quantum gravity*, JHEP 0207 (2002) 062 [hep-th/0202018]
22. J. Ambjørn and C.F. Kristjansen: *Nonperturbative 2-d quantum gravity and Hamiltonians unbounded from below*, Int. J. Mod. Phys. A 8 (1993) 1259–1282 [hep-th/9205073]

23. L. Freidel and D. Louapre: *Non-perturbative summation over 3D discrete topologies* [hep-th/0211026]
24. J. Ambjørn, J. Jurkiewicz and R. Loll: *Nonperturbative 3d Lorentzian quantum gravity*, Phys. Rev. D 64 (2001) 044011 [hep-th/0011276]
25. J. Ambjørn, J. Jurkiewicz and R. Loll: *Computer simulations of 3d Lorentzian quantum gravity*, Nucl. Phys. B (Proc. Suppl.) 94 (2001) 689–692 [hep-lat/0011055].
26. P. Di Francesco, E. Guitter and C. Kristjansen: *Integrable 2-d Lorentzian gravity and random walks*, Nucl. Phys. B 567 (2000) 515–553 [hep-th/9907084]
27. P. Di Francesco, E. Guitter and C. Kristjansen: *Generalized Lorentzian triangulations and the Calogero Hamiltonian*, Nucl. Phys. B 608 (2001) 485–526 [hep-th/0010259]
28. J. Ambjørn, R. Loll, J.L. Nielsen and J. Rolf: Euclidean and Lorentzian quantum gravity – lessons from two dimensions, *J. Chaos Solitons Fractals* 10 (1999) 177–195 [hep-th/9806241]
29. J. Ambjørn, J. Correia, C. Kristjansen and R. Loll: *The relation between Euclidean and Lorentzian 2D quantum gravity*, Phys. Lett. B 475 (2000) 24–32 [hep-th/9912267]
30. H. Aoki, H. Kawai, J. Nishimura and A. Tsuchiya: *Operator product expansion in two-dimensional quantum gravity*, Nucl. Phys. B 474 (1996) 512–528 [hep-th/9511117]
31. D.V. Boulatov and V.A. Kazakov: *The Ising model on random planar lattice: the structure of phase transition and the exact critical exponents*, Phys. Lett. B 186B (1987) 379
32. J. Ambjørn, K.N. Anagnostopoulos and R. Loll: *A new perspective on matter coupling in 2d quantum gravity*, Phys. Rev. D 60 (1999) 104035 [hep-th/9904012]
33. J. Ambjørn, K.N. Anagnostopoulos and R. Loll: *Crossing the c=1 barrier in 2d Lorentzian quantum gravity*, Phys. Rev. D 61 (2000) 044010 [hep-lat/9909129]
34. J. Ambjørn, J. Jurkiewicz and R. Loll: *3d Lorentzian, dynamically triangulated quantum gravity*, Nucl. Phys. Proc. Suppl. 106 (2002) 980–982 [hep-lat/0201013]
35. J. Ambjørn, A. Dasgupta, J. Jurkiewicz and R. Loll: *A Lorentzian cure for Euclidean troubles*, Nucl. Phys. Proc. Suppl. 106 (2002) 977–979 [hep-th/0201104]
36. J. Ambjørn, J. Jurkiewicz, R. Loll and G. Vernizzi: *Lorentzian 3d gravity with wormholes via matrix models*, JHEP 0109 (2001) 022 [hep-th/0106082]
37. V.A. Kazakov and P. Zinn-Justin: *Two matrix model with ABAB interaction*, Nucl. Phys. B 546 (1999) 647–668 [hep-th/9808043]
38. J. Ambjørn, R. Janik, J. Jurkiewicz, R. Loll and G. Vernizzi, work in progress
39. C. Dehne: *Konstruktionsversuche eines quantenkosmologischen, dynamisch triangulierten Torusuniversums in 2+1 Dimensionen* (in German), Diploma Thesis, Univ. Hamburg (2001), http://www.aei-potsdam.mpg.de/research/thesis/dehne_dipl.ps.gz
40. B. Dittrich and R. Loll: *A hexagon model for 3D Lorentzian quantum cosmology*, Phys. Rev. D 66 (2002) 084016 [hep-th/0204210]
41. R. Costa-Santos and R. Loll, work in progress
42. D. Kappel: *Nichtperturbative Pfadintegrale der Quantengravitation durch kausale dynamische Triangulierungen* (in German), Diploma Thesis, Univ. Potsdam (2001)
43. B. Dittrich: *Dynamische Triangulierung von Schwarzloch-Geometrien* (in German), Diploma Thesis, Univ. Potsdam (2001)
44. C. Rovelli and L. Smolin: *Discreteness of area and volume in quantum gravity*, Nucl. Phys. B 442 (1995) 593–622, Erratum-ibid. B 456 (1995) 753 [gr-qc/9411005]

45. R. Loll: *The volume operator in discretized quantum gravity*, Phys. Rev. Lett. 75 (1995) 3048–3051 [gr-qc/9506014].
46. R. Loll: *Spectrum of the volume operator in quantum gravity*, Nucl. Phys. B 460 (1996) 143–154 [gr-qc/9511030].

Introduction to String Theory

Thomas Mohaupt

Friedrich-Schiller Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany

Abstract. We give a pedagogical introduction to string theory, D-branes and p-brane solutions.

1 Introductory Remarks

These notes are based on lectures given at the 271-th WE-Heraeus-Seminar 'Aspects of Quantum Gravity'. Their aim is to give an introduction to string theory for students and interested researches. No previous knowledge of string theory is assumed. The focus is on gravitational aspects and we explain in some detail how gravity is described in string theory in terms of the graviton excitation of the string and through background gravitational fields. We include Dirichlet boundary conditions and D-branes from the beginning and devote one section to p-brane solutions and their relation to D-branes. In the final section we briefly indicate how string theory fits into the larger picture of M-theory and mention some of the more recent developments, like brane world scenarios.

The WE-Heraeus-Seminar 'Aspects of Quantum Gravity' covered both main approaches to quantum gravity: string theory and canonical quantum gravity. Both are complementary in many respects. While the canonical approach stresses background independence and provides a non-perturbative framework, the cornerstone of string theory still is perturbation theory in a fixed background geometry. Another difference is that in the canonical approach gravity and other interactions are independent from each other, while string theory automatically is a unified theory of gravity, other interactions and matter. There is a single dimensionful constant and all couplings are functions of this constant and of vacuum expectation values of scalars. The matter content is uniquely fixed by the symmetries of the underlying string theory. Moreover, when formulating the theory in Minkowski space, the number of space-time dimensions is fixed. As we will see, there are only five distinct supersymmetric string theories in ten-dimensional Minkowski space.

The most important feature of string perturbation theory is the absence of UV divergencies. This allows one to compute quantum corrections to scattering amplitudes and to the effective action, including gravitational effects. More recently, significant progress has been made in understanding non-perturbative aspects of the theory, through the study of solitons and instantons, and through string dualities which map the strong coupling behaviour of one string theory to the weak coupling behaviour of a dual theory. Moreover, string dualities relate

T. Mohaupt, Introduction to String Theory, Lect. Notes Phys. **631**, 173–251 (2003)
http://www.springerlink.com/

all five supersymmetric string theories to one another and lead to the picture of one single underlying theory, called M-theory. So far, only various limits of this theory are known, while the problem of finding an intrinsic, non-perturbative and background-independent definition is unsolved. One expects that M-theory has an underlying principle which unifies its various incarnations, presumably a symmetry principle. One of the obstacles on the way to the final theory is that it is not clear which degrees of freedom are fundamental. Besides strings, also higher-dimensional p-branes play an essential role. Moreover, there is an eleven-dimensional limit, which cannot be described in terms of strings.

Our presentation of string theory will be systematic rather than follow the path of historical development. Nevertheless we feel that a short historical note will be helpful, since many aspects which may seem somewhat ad hoc (such as the definition of interactions in Sect. 3) become clearer in their historical context. The story started with the Veneziano amplitude, which was proposed as an amplitude for meson scattering in pre-QCD times. The amplitude fitted the known experimental data very well and had precisely the properties expected of a good scattering amplitude on the basis of S-matrix theory, the bootstrap program and Regge pole theory. In particular it had a very special soft UV behaviour. Later work by Y. Nambu, H.B. Nielsen and L. Susskind showed that the Veneziano amplitude, and various generalizations thereof could be interpreted as describing the scattering of relativistic strings. But improved experimental data ruled out the Veneziano amplitude as a hadronic amplitude: it behaved just to softly in order to describe the hard, partonic substructures of hadrons seen in deep inelastic scattering. J. Scherk and J. Schwarz reinterpreted string theory as a unified theory of gravity and all other fundamental interactions, making use of the fact that the spectrum of a closed string always contains a massless symmetric tensor state which couples like a graviton. This led to the development of perturbative string theory, as we will describe it in Sects. 2–4 of these lecture notes. More recently the perspective has changed again, after the role of D-branes, p-branes and string dualities was recognized. This will be discussed briefly in Sects. 5 and 6.

From the historical perspective it appears that string theory is a theory which is 'discovered' rather than 'invented'. Though it was clear from the start that one was dealing with an interesting generalization of quantum field theory and general relativity, the subject has gone through several 'phase transitions', and its fundamental principles remain to be made explicit. This is again complementary to canonical quantum gravity, where the approach is more axiomatic, starting from a set of principles and proceeding to quantize Einstein gravity.

The numerous historical twists, our lack of final knowledge about the fundamental principles and the resulting diversity of methods and approaches make string theory a subject which is not easy to learn (or to teach). The 271-th WE-Heraeus-Seminar covered a broad variety of topics in quantum gravity, 'From Theory to Experimental Search'. The audience consisted of two groups: graduate students, mostly without prior knowledge of string theory, and researchers, working on various theoretical and experimental topics in gravity. The two lectures on string theory were supposed to give a pedagogical introduction and to prepare for later lectures on branes worlds, large extra dimensions, the AdS-CFT correspondence and black holes. These lecture notes mostly follow the lectures, but aim to extend them in two ways. The first is to add more details to the topics I discussed in the lectures. In particular I want to expand on points which seemed to be either difficult or interesting to the audience. The second goal is to include more material, in order to bring the reader closer to the areas of current active research. Both goals are somewhat contradictory, given that the result is not meant to be a book, but lecture notes of digestable length. As a compromise I choose to explain those things in detail which seemed to be the most important ones for the participants of the seminar, hoping that they represent a reasonable sample of potential readers. On the other side several other topics are also covered, though in a more scetchy way. Besides summarizing advanced topics, which cannot be fully explained here, I try to give an overview of (almost) all the new developements of the last years and to indicate how they fit into the emerging overall picture of M-theory.

The outline of the lectures is as follows: Sects. 2–4 are devoted to perturbative aspects of bosonic and supersymmetric string theories. They are the core of the lectures. References are given at the end of the sections. String theory has been a very active field over several decades, and the vast amount of existing literature is difficult to oversee even for people working in the field. I will not try to give a complete account of the literature, but only make suggestions for further reading. The basic references are the books [1–5], which contain a huge number of references to reviews and original papers. The reader interested in the historical developement of the subject will find information in the annotated bibliography of [1]. Section 5 gives an introduction to non-perturbative aspects by discussing a particular class of solitons, the p-brane solutions of type II string theory. Section 6 gives an outlook on advanced topics: while Sects. 6.1–6.3 scetch how the five supersymmetric string theories fit into the larger picture of M-theory, Sect. 6.4 gives an overview of current areas of research, together with references to lecture notes, reviews and some original papers.

2 Free Bosonic Strings

We start our study of string theories with the bosonic string. This theory is a toy-model rather than a realistic theory of gravity and matter. As indicated by its name it does not have fermionic states, and this disqualifies it as a theory of particle physics. Moreover, its ground state in Minkowski space is a tachyon, i.e., a state of negative mass squared. This signals that the theory is unstable. Despite these shortcomings, the bosonic string has its virtues as a pedagogical toy-model: whereas we can postpone to deal with the additional techniques needed to describe fermions, many features of the bosonic string carry over to supersymmetric string theories, which have fermions but no tachyon.

2.1 Classical Bosonic Strings

We start with a brief overview of classical aspects of bosonic strings.

Setting the Stage. Let us first fix our notation. We consider a fixed background Pseudo-Riemannian space-time $\mathcal{M}$ of dimension D, with coordinates $X = (X^\mu)$, $\mu = 0, \ldots, D-1$. The metric is $G_{\mu\nu}(X)$ and we take the signature to be 'mostly plus', $(-)(+)^{D-1}$.

The motion of a relativistic string in $\mathcal{M}$ is described by its generalized world-line, a two-dimensional surface Σ, which is called the world-sheet. For a single non-interacting string the world-sheet has the form of an infinite strip. We introduce coordinates $\sigma = (\sigma^0, \sigma^1)$ on the world-sheet. The embedding of the world-sheet into space-time is given by maps

$$X : \Sigma \longrightarrow \mathcal{M} \; : \sigma \longrightarrow X(\sigma) \; . \tag{1}$$

The background metric induces a metric on the world-sheet:

$$G_{\alpha\beta} = \frac{\partial X^\mu}{\partial \sigma^\alpha} \frac{\partial X^\nu}{\partial \sigma^\beta} G_{\mu\nu} \; , \tag{2}$$

where $\alpha, \beta = 0, 1$ are world-sheet indices. The induced metric is to be distinguished from an intrinsic metric $h_{\alpha\beta}$ on Σ. As we will see below, an intrinsic metric is used as an auxiliary field in the Polyakov formulation of the bosonic string.

A useful, but sometimes confusing fact is that the above setting can be viewed from two perspectives. So far we have taken the space-time perspective, interpreting the system as a relativistic string moving in space-time $\mathcal{M}$. Alternatively we may view it as a two-dimensional field theory living on the world-sheet, with fields X which take values in the target-space $\mathcal{M}$. This is the world-sheet perspective, which enables us to use intuitions and methods of two-dimensional field theory for the study of strings.

Actions. The natural action for a relativistic string is its area, measured with the induced metric:

$$S_{\mathrm{NG}} = \frac{1}{2\pi\alpha'} \int_\Sigma d^2\sigma |\det G_{\alpha\beta}|^{1/2} \; . \tag{3}$$

This is the Nambu-Goto action, which is the direct generalization of the action for a massive relativistic particle. The prefactor $(2\pi\alpha')^{-1}$ is the energy per length or tension of the string, which is the fundamental dimensionful parameter of the theory. We have expressed the tension in terms of the so-called Regge slope α', which has the dimension $(\text{length})^2$ in natural units, $c = 1$, $\hbar = 1$. Most of the time we will use string units, where in addition we set $\alpha' = \frac{1}{2}$.

The Nambu-Goto action has a direct geometric meaning, but is technically inconvenient, due to the square root. Therefore one prefers to use the Polyakov action, which is equivalent to the Nambu-Goto action, but is a standard two-dimensional field theory action. In this approach one introduces an intrinsic metric on the world-sheet, $h_{\alpha\beta}(\sigma)$, as additional datum. The action takes the form of a non-linear sigma-model on the world-sheet,

$$S_{\mathrm{P}} = \frac{1}{4\pi\alpha'} \int_{\Sigma} d^2\sigma \sqrt{h} h^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X^\nu G_{\mu\nu}(X) , \tag{4}$$

where $h = |\det h_{\alpha\beta}|$.

The equation of motion for $h_{\alpha\beta}$ is algebraic. Thus the intrinsic metric is non-dynamical and can be eliminated, which brings us back to the Nambu-Goto action. Since

$$T_{\alpha\beta} := \left(2\pi\alpha'\sqrt{h}\right)^{-1} \frac{\delta S_{\mathrm{P}}}{\delta h^{\alpha\beta}} = \partial_\alpha X^\mu \partial_\beta X_\mu - \frac{1}{2} h_{\alpha\beta} \partial_\gamma X^\mu \partial^\gamma X_\mu \tag{5}$$

is the energy momentum of the two-dimensional field theory defined by (4), we can interpret the equation of motion of $h_{\alpha\beta}$ as the two-dimensional Einstein equation. The two-dimensional metric is non-dynamical, because the two-dimensional Einstein-Hilbert action is a topological invariant, proportional to the Euler number of Σ. Thus its variation vanishes and the Einstein equation of (4) coupled to two-dimensional gravity reduces to $T_{\alpha\beta} = 0$. Note that the energy-momentum tensor (5) is traceless, $h^{\alpha\beta} T_{\alpha\beta} = 0$. This holds before imposing the equations of motion ('off shell'). Therefore $T_{\alpha\beta}$ has only two independent components, which vanish for solutions to the equations of motion ('on shell'). Since the trace of the energy-momentum tensor is the Noether current of scale transformations, this shows that the two-dimensional field theory (4) is scale invariant. As we will see below, it is in fact a conformal field theory.

The Polyakov action has three local symmetries. Two are shared by the Nambu-Goto action, namely reparametrizations of the world-sheet:

$$\sigma^\alpha \longrightarrow \tilde{\sigma}^\alpha(\sigma^0, \sigma^1) . \tag{6}$$

The third local symmetry is the multiplication of the metric $h_{\alpha\beta}$ by a local, positive scale factor,

$$h_{\alpha\beta}(\sigma) \longrightarrow e^{\Lambda(\sigma)} h_{\alpha\beta}(\sigma) . \tag{7}$$

This transformation is called a Weyl transformation by physicists, while mathematicians usually use the term conformal transformation. The three local symmetries can be used to gauge-fix the metric $h_{\alpha\beta}$. The standard choice is the conformal gauge,

$$h_{\alpha\beta}(\sigma) \stackrel{!}{=} \eta_{\alpha\beta} , \quad \text{where} \quad (\eta_{\alpha\beta}) = \mathrm{Diag}(-1, 1) . \tag{8}$$

While this gauge can be imposed globally on the infinite strip describing the motion of a single non-interacting string, it can only be imposed locally on more general world-sheets, which describe string interactions. We will discuss global aspects of gauge fixing later.

The conformal gauge does not provide a complete gauge fixing, because (8) is invariant under a residual symmetry. One can still perform reparametrizations under which the metric only changes by a local, positive scale factor, because this factor can be absorbed by a Weyl transformation. Such conformal reparametrizations are usually called conformal transformations by physicists. Note that the same term is used for Weyl transformations by mathematicians. A convenient way to characterize conformal reparametrizations in terms of coordinates is to introduce light cone coordinates,

$$\sigma^{\pm} = \sigma^0 \pm \sigma^1 \,. \tag{9}$$

Then conformal reparametrization are precisely those reparametrizations which do not mix the light cone coordinates:

$$\sigma^+ \longrightarrow \tilde{\sigma}^+(\sigma^+) \,, \quad \sigma^- \longrightarrow \tilde{\sigma}^-(\sigma^-) \,. \tag{10}$$

Thus we are left with an infinite-dimensional group of symmetries, which in particular includes scale transformations.

Equations of Motion, Closed and Open Strings, and D-Branes. In order to proceed we now specialize to the case of a flat space-time, $G_{\mu\nu} = \eta_{\mu\nu}$, where $\eta_{\mu\nu} = \mathrm{Diag}(-1, +1, \ldots, +1)$. In the conformal gauge the equation of motion for X reduces to a free two-dimensional wave equation,

$$\partial^2 X^\mu = \partial^\alpha \partial_\alpha X^\mu = 0 \,. \tag{11}$$

Note that when imposing the conformal gauge on the Polyakov action (4), the equation of motion for $h_{\alpha\beta}$, i.e. $T_{\alpha\beta} = 0$, becomes a constraint, which has to be imposed on the solutions of (11).

The general solution of (11) is a superposition of left- and right-moving waves,

$$X^\mu(\sigma) = X^\mu_L(\sigma^+) + X^\mu_R(\sigma^-) \,. \tag{12}$$

However, we also have to specify boundary conditions at the ends of the string. One possible choice are periodic boundary conditions,

$$X^\mu(\sigma^0, \sigma^1 + \pi) = X^\mu(\sigma^0, \sigma^1) \,. \tag{13}$$

They correspond to closed strings. A convenient parametrization of the solution is:

$$X^\mu(\sigma) = x^\mu + 2\alpha' p^\mu \sigma^0 + \mathrm{i}\sqrt{2\alpha'} \sum_{n\neq 0} \frac{\alpha^\mu_n}{n} e^{-2\mathrm{i}n\sigma^+} + \mathrm{i}\sqrt{2\alpha'} \sum_{n\neq 0} \frac{\tilde{\alpha}^\mu_n}{n} e^{-2\mathrm{i}n\sigma^-} \,. \tag{14}$$

Reality of X^μ implies: $(x^\mu)^\star = x^\mu$ and $(p^\mu)^\star = p^\mu$ and $(\alpha^\mu_m)^\star = \alpha^\mu_{-m}$ and $(\tilde{\alpha}^\mu_m)^\star = \tilde{\alpha}^\mu_{-m}$. Here $\star$ denotes complex conjugation. While x^μ is the position of the center of mass of the string at time σ^0, p^μ is its total momentum. Thus, the center of mass moves on a straight line in Minkowski space, like a free relativistic particle. The additional degrees of freedom are decoupled left- and right-moving waves on the string, with Fourier components α_m^μ and $\tilde{\alpha}_m^\mu$.

When not choosing periodic boundary conditions, the world-sheet has boundaries and we have open strings. The variation of the world-sheet action yields a boundary term, $\delta S \simeq \int_{\partial\Sigma} d\sigma^0 \partial_1 X^\mu \delta X_\mu$. The natural choice to make the boundary term vanish are Neumann boundary conditions,

$$\partial_1 X^\mu|_{\sigma^1=0} = 0 \;, \quad \partial_1 X^\mu|_{\sigma^1=\pi} = 0 \;. \tag{15}$$

With these boundary conditions, momentum is conserved at the ends of the string. Left- and right-moving waves are reflected at the ends and combine into standing waves. The solution takes the form

$$X^\mu(\sigma) = x^\mu + 2\alpha' p^\mu \sigma^0 + \mathrm{i}\sqrt{2\alpha'} \sum_{n\neq 0} \frac{\alpha_n^\mu}{n} e^{-\mathrm{i}n\sigma^0} \cos(n\sigma^1) \;. \tag{16}$$

There is, however, a second possible choice of boundary conditions for open strings, namely Dirichlet boundary conditions. Here the ends of the string are kept fixed:

$$X^\mu|_{\sigma^1=0} = x^\mu_{(1)} \;, \quad X^\mu|_{\sigma^1=\pi} = x^\mu_{(2)} \;. \tag{17}$$

With these boundary conditions the solution takes the form

$$X^\mu(\sigma) = x^\mu_{(1)} + (x^\mu_{(2)} - x^\mu_{(1)})\frac{\sigma^1}{\pi} + \mathrm{i}\sqrt{2\alpha'} \sum_{n\neq 0} \frac{\alpha_n^\mu}{n} e^{-\mathrm{i}n\sigma^0} \sin(n\sigma^1) \;. \tag{18}$$

More generally we can impose Neumann boundary conditions in the time and in p space directions and Dirichlet boundary conditions in the other directions. Let us denote the Neumann directions by $(X^m) = (X^0, X^1, \ldots, X^p)$ and the Dirichlet directions by $(X^a) = (X^{p+1}, \ldots, X^{D-1})$.

The most simple choice of Dirichlet boundary conditions is then to require that all open strings begin and end on a p-dimensional plane located at an arbitrary position $X^a = x^a_{(1)}$ along the Dirichlet directions. Such a plane is called a p-dimensional Dirichlet-membrane, or D-p-brane, or simply D-brane for short. While the ends of the strings are fixed in the Dirichlet directions, they still can move freely along the Neumann directions. The world-volume of a D-p-brane is (p + 1)-dimensional. The Neumann directions are called the world-volume or the parallel directions, while the Dirichlet directions are called transverse directions.

An obvious generalization is to introduce $N > 1$ such D-p-branes, located at positions $x^a_{(i)}$, where $i = 1, \ldots, N$, and to allow strings to begin and end on any of these. In this setting the mode expansion for a string starting on the i-th D-brane and ending on the j-th is:

$$\begin{aligned} X^m(\sigma) &= x^m + 2\alpha' p^m \sigma^0 + \mathrm{i}\sqrt{2\alpha'} \sum_{n\neq 0} \frac{\alpha_n^m}{n} e^{-\mathrm{i}n\sigma^0} \cos(n\sigma^1) \;, \\ X^a(\sigma) &= x^a_{(i)} + (x^a_{(j)} - x^a_{(i)})\frac{\sigma^1}{\pi} + \mathrm{i}\sqrt{2\alpha'} \sum_{n\neq 0} \frac{\alpha_n^a}{n} e^{-\mathrm{i}n\sigma^0} \sin(n\sigma^1) \;. \end{aligned} \tag{19}$$

(One might also wonder about Dirichlet boundary conditions in the time direction. This makes sense, at least for Euclidean space-time signature, and leads to instantons, called D-instantons, which we will not discuss in these lectures.)

Dirichlet boundary conditions have been neglected for several years. The reason is that momentum is not conserved at the ends of the strings, reflecting that translation invariance is broken along the Dirichlet directions. Therefore, in a complete fundamental theory the D-branes must be new dynamical objects, different from strings. The relevance of such objects was only appreciated when it became apparent that string theory already includes solitonic space-time backgrounds, so called ('RR-charged') p-Branes, which correspond to D-branes. We will return to this point later.

Promoting the D-branes to dynamical objects implies that they will interact through the exchange of strings. This means that in general they will repulse or attract, and therefore their positions become dynamical. But there exist many static configurations of D-branes (mainly in supersymmetric string theories), where the attractive and repulsive forces cancel for arbitrary distances of the branes.

2.2 Quantized Bosonic Strings

The definition of a quantum theory of bosonic strings proceeds by using standard recipies of quantization. The two most simple ways to proceed are called 'old covariant quantization' and 'light cone quantization'. As mentioned above, imposing the conformal gauge leaves us with a residual gauge invariance. In light cone quantization one fixes this residual invariance by imposing the additional condition

$$X^+ \stackrel{!}{=} x^+ + p^+\sigma^+ \;, \quad \text{i.e.}\;, \;\; \alpha_m^+ \stackrel{!}{=} 0 \;, \tag{20}$$

where $X^\pm = \frac{1}{\sqrt{2}}(X^0 \pm X^{D-1})$ are light cone coordinates in space-time. Then the constraints $T_{\alpha\beta} = 0$ are solved in the classical theory. This yields (non-linear) expressions for the oscillators α_n^- in terms of the transverse oscillators α_n^i, $i = 1, \ldots D-2$. In light cone coordinates the world-sheet is embedded into space-time along the X^0, X^{D-1} directions. The independent degrees of freedom are the oscillations transverse to the world sheet, which are parametrized by the α_n^i. One proceeds to quantize these degrees of freedom. In this approach unitarity of the theory is manifest, but Lorentz invariance is not.

In old covariant quantization one imposes the constraints at the quantum level. Lorentz covariance is manifest, but unitarity is not: one has to show that there is a positive definite space of states and a unitary S-matrix. This is the approach we will describe in more detail below.

One might also wonder about 'new covariant quantization', which is BRST quantization. This approach is more involved but also more powerful than old covariant quantization. When dealing with advanced technical problems, for example the construction of scattering amplitudes involving fermions in superstring theories, BRST techniques become mandatory. But this is beyond the scope of these lectures.

The Fock Space. The first step is to impose canonical commutation relations on $X^\mu(\sigma)$ and its canonical momentum $\Pi^\mu(\sigma) = \partial_0 X^\mu(\sigma)$. In terms of modes one gets

$$[x^\mu, p^\nu] = \mathrm{i}\eta^{\mu\nu} \ , \quad [\alpha^\mu_m, \alpha^\nu_n] = m\eta^{\mu\nu}\delta_{m+n,0} \ . \tag{21}$$

For closed strings there are analogous relations for $\tilde{\alpha}^\mu_m$. The reality conditions of the classical theory translate into hermiticity relations:

$$(x^\mu)^\dagger = x^\mu \ , \quad (p^\mu)^\dagger = p^\mu \ , \quad (\alpha^\mu_m)^\dagger = \alpha^\mu_{-m} \ . \tag{22}$$

While the commutation relations for x^μ, p^ν are those of a relativistic particle, the α^μ_m satisfy the relations of creation and annihilation operators of harmonic oscillators, though with an unconventional normalization.

To proceed, one constructs a Fock space $\mathcal{F}$ on which the commutation relations (21) are represented. First one chooses momentum eigenstates $|k\rangle$, which are annihiliated by half of the oscillators:

$$p^\mu|k\rangle = k^\mu|k\rangle \ , \quad \alpha^\mu_m|k\rangle = 0 = \tilde{\alpha}^\mu_m|k\rangle \ , \quad m > 0 \ . \tag{23}$$

Then a basis $\mathcal{B}$ of $\mathcal{F}$ is obtained by acting with creation operators:

$$\mathcal{B} = \{\alpha^{\mu_1}_{-m_1} \cdots \tilde{\alpha}^{\nu_1}_{-n_1} \cdots |k\rangle \mid m_l, n_l > 0\} \ . \tag{24}$$

A bilinear form on $\mathcal{F}$ which is compatible with the hermiticity properties (22) cannot be positive definite. Consider for example the norm squared of the state $\alpha^\mu_{-m}|k\rangle$:

$$\langle k|(\alpha^\mu_{-m})^+\alpha^\mu_{-m}|k\rangle \sim \eta^{\mu\mu} = \pm 1 \ . \tag{25}$$

However, the Fock space is not the space of physical states, because we still have to impose the constraints. The real question is whether the subspace of physical states contains states of negative norm.

The Virasoro Algebra. Constraints arise when the canonical momenta of a system are not independent. This is quite generic for relativistic theories. The most simple example is the relativistic particle, where the constraint is the mass shell condition, $p^2 + m^2 = 0$. When quantizing the relativistic particle, physical states are those annihilated by the constraint, i.e., states satisfying the mass shell condition:

$$(p^2 + m^2)|\Phi\rangle = 0 \ . \tag{26}$$

When evaluating this in a basis of formal eigenstates of the operator x^μ, one obtains the Klein-Gordon equation, $(\partial^2 + m^2)\Phi(x) = 0$, where $\Phi(x) = \langle x|\Phi\rangle$ is interpreted as the state vector in the x-basis. This is a clumsy way to approach the quantum theory of relativistic particles, and one usually prefers to use quantum field theory ('second quantization') rather than quantum mechanics ('first quantization'). But in string theory it turns out that the first quantized formulation works nicely for studying the spectrum and computing amplitudes, whereas string field theory is very complicated.

Proceeding parallel to the case of a relativistic particle one finds that the canonical momentum is $\Pi^\mu = \partial_0 X^\mu$. The constraints are

$$\Pi^\mu \partial_1 X_\mu = 0 \;, \quad \Pi^\mu \Pi_\mu + \partial_1 X^\mu \partial_1 X_\mu = 0 \;. \tag{27}$$

In the Polyakov formulation they are equivalent to $T_{\alpha\beta} = 0$. It is convenient to express the constraints through the Fourier components of $T_{\alpha\beta}$. Passing to light cone coordinates, the tracelessness of $T_{\alpha\beta}$, which holds without using the equation of motion or imposing the constraints, implies

$$T_{+-} = 0 = T_{-+} \;. \tag{28}$$

Thus we are left with two independent components, T_{++} and T_{--}, where $T_{\pm\pm} \simeq \partial_\pm X^\mu \partial_\pm X_\mu$. For closed strings, where $\partial_\pm X^\mu$ are periodic in σ^1, we expand $T_{\pm\pm}$ in a Fourier series and obtain Fourier coefficients $L_m, \tilde{L}_m$, $m \in \mathbf{Z}$. For open strings, observe that $\sigma^1 \to -\sigma^1$ exchanges $\partial_+ X^\mu$ and $\partial_- X^\mu$. Both fields can be combined into a single field, which is periodic on a formally doubled world-sheet with $-\pi \leq \sigma^1 \leq \pi$. In the same way one can combine T_{++} with T_{--}. By Fourier expansion on the doubled world-sheet one then obtains one set of Fourier modes for the energy-momentum tensor, denoted L_m. This reflects that left- and right-moving waves couple through the boundaries.

The explicit form for the L_m in terms of oscillators is

$$L_m = \frac{1}{2} \sum_{n=-\infty}^{\infty} \alpha_{m-n} \cdot \alpha_n \;, \tag{29}$$

with an analogous formula for $\tilde{L}_m$ for closed strings. We have denoted the contraction of Lorentz indices by '$\cdot$' and defined $\alpha_0^\mu = \frac{1}{2} p^\mu = \tilde{\alpha}_0^\mu$ for closed strings and $\alpha_0^\mu = p^\mu$ for open strings. In terms of the Fourier modes, the constraints are $L_m = 0$ and, for closed strings, $\tilde{L}_m = 0$. Translations in σ^0 are generated by L_0 for open and by $L_0 + \tilde{L}_0$ for closed strings. These functions are the world-sheet Hamiltonians. The L_m satisfy the Witt algebra,

$$\{L_m, L_n\}_{\text{P.B.}} = \mathrm{i}(m-n) L_{m+n} \;, \tag{30}$$

where $\{\cdot,\cdot\}_{\text{P.B.}}$ is the Poisson bracket. For closed strings we have two copies of this algebra. The Witt algebra is the Lie algebra of infinitesimal conformal transformations. Thus the constraints reflect that we have a residual gauge symmetry corresponding to conformal transformations. Since the constraints form a closed algebra with the Hamiltonian, they are preserved in time. Such constraints are called first class, and they can be imposed on the quantum theory without further modifications (such as Dirac brackets).

In the quantum theory the L_m are taken to be normal ordered, i.e., annihilation operators are moved to the right. This is unambigous, except for L_0. We will deal with this ordering ambiguity below. The hermiticiy properties of the L_m are:

$$L_m^\dagger = L_{-m} \ . \tag{31}$$

The operators L_m satisfy the Virasora algebra:

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}(m^3 - m)\delta_{m+n,0} \ . \tag{32}$$

The Virasoro algebra is a central extension of the Witt algebra. On our Fock space $\mathcal{F}$ the central charge c takes the value

$$c = \eta^{\mu\nu}\eta_{\mu\nu} = D \ , \tag{33}$$

i.e., each space-time dimension contributes one unit. Since the Poisson brackets of L_m in the classical theory just give the Witt algebra, this dependence on the number of dimensions is a new property of the quantum theory. The extra central term occuring at the quantum level is related to a normal ordering ambiguity of commutators with $m + n = 0$. This results in a new 'anomalous' term in the algebra. In the context of current algebras such terms are known as Schwinger terms.

Imposing the Constraints, or, Why $D = 26$? In the classical theory the constraints amount to imposing $L_m = 0$ on solutions. Imposing this as an operator equation on the quantum theory is too strong. In particular it is not compatible with the algebra (32). What can be imposed consistently is that matrix elements of the L_m vanish between physical states, $\langle \Phi_1 | L_m | \Phi_2 \rangle = 0$. Conversely this condition singles out the subspace of physical states, $\mathcal{F}_{\text{phys}} \subset \mathcal{F}$. Using the hermiticity properties of the L_m, this is equivalent to the statement that the positive Virasoro modes annihilate physical states,

$$\begin{aligned} L_m|\Phi\rangle &= 0 \ , \quad m > 0 \ , \\ (L_0 - a)|\Phi\rangle &= 0 \ , \end{aligned} \tag{34}$$

for all $|\Phi\rangle \in \mathcal{F}_{\text{phys}}$. Note that we have introduced an undetermined constant a into the L_0-constraint. As mentioned above this operator has an ordering ambiguity. We take L_0 to be normal ordered and parametrize possible finite ordering effects by the constant a. Since L_0 is the Hamiltonian, this might be considered as taking into account a non-trivial Casimir effect. In the case of closed strings there is a second set of constraints involving the $\tilde{L}_m$.

The Virasoro operators L_{-m}, $m > 0$ still act non-trivially on physical states and create highest weight representations of the Virasoro algebra. This corresponds to the fact that we still have residual gauge symmetries. Therefore it is clear that $\mathcal{F}_{\text{phys}}$ is not the physical Hilbert space. $\mathcal{F}_{\text{phys}}$ is not positive definite, but contains null states (states of norm zero) and, depending on the number of space-time dimensions, also states of negative norm. A positive definite space of states can be constructed if negative norm states are absent, such that $\mathcal{F}_{\mathrm{phys}}$ is positive semi-definite, and if null states are orthogonal to all physical states. Then one can consistently identify physical states $|\Phi\rangle$ that differ by null states $|\Psi\rangle$,

$$|\Phi\rangle \simeq |\Phi\rangle + |\Psi\rangle \;, \tag{35}$$

and define the Hilbert space by

$$\mathcal{H} = \mathcal{F}_{\mathrm{phys}}/\{\text{Null states}\} \;. \tag{36}$$

The working of this construction crucially depends on the values of D and a. This is the contents of the so-called no-ghost theorem, which can be summarized as follows:

1. $D = 26$ and $a = 1$. The construction works as described above. The resulting theory is known as the critical (bosonic) string theory, $D = 26$ is the critical dimension. Physical states differing by a null states differ by a residual gauge transformation and represent the same state in the Hilbert space. We will see explicit examples below.
2. $D > 26$. The physical subspace $\mathcal{F}_{\mathrm{phys}}$ always contains states of negative norm and no Hilbert space $\mathcal{H}$ can be constructed. There is no bosonic string theory for $D > 26$.
3. $D \leq 25$. Naively one expects such theories to be unitary, because we can just truncate the unitary critical string theory and this cannot introduce states of negative norm. Nevertheless one does not obtain a consistent quantum theory by truncation. When studying scattering amplitudes at the loop level one finds poles corresponding to unphysical negative norm states and there is no unitary S-matrix. Thus truncations of the critical string do not yield unitary theories.
   But there is an alternative to truncation, known as Liouville string theory or non-critical string theory. This theory exists in $D < 26$, at the price that the quantum theory has a new degree of freedom, the Liouville mode. (This is most obvious in a path integral formulation.) The resulting theory is much more complicated than the critical string, because its world-sheet theory is interacting even for a flat target space. For this theory much less is known than about the critical string. However, there are arguments indicating that the non-critical string is equivalent to the critical string in a non-trivial background.

We will only consider critical string theories in the following. Also note that the above analysis applies to strings in flat space-time, with no background fields. When switching on a non-trivial dilaton background, this can modify the central charge of the world-sheet conformal field theory, and, hence, the dimension of space-time. But this topic is beyond the scope of these lectures.

The Spectrum of the Bosonic Closed String. We can now identify the physical states by imposing the constraints. Let us consider closed strings. We first look at the two constraints

$$(L_0 - 1)|\Phi\rangle = 0 \;, \quad (\tilde{L}_0 - 1)|\Phi\rangle = 0 \;. \tag{37}$$

The operator L_0 can be rewritten as

$$L_0 = \tfrac{1}{8}p^2 + N \;. \tag{38}$$

As mentioned above the operator L_0 is the normal ordered version of (29) with $m = 0$. The original and the normal ordered expression formally differ by an infinite constant. Subtracting this constant introduces a finite ambiguity, which was parametrized by a. Unitarity then fixes $a = 1$. The oscillator part of L_0 is

$$N = \sum_{n=1}^{\infty} \alpha_{-n} \cdot \alpha_n \;. \tag{39}$$

N is called the number operator, because

$$[N, \alpha^\mu_{-m}] = m\alpha^\mu_{-m} \;. \tag{40}$$

Since the total momentum is related to the mass of the string by $M^2 + p^2 = 0$, the constraints (37) determine the mass of a physical states in terms of the eigenvalues of N and of its right-moving analogue $\tilde{N}$. (We denote the operators and their eigenvalues by the same symbol.) We now use the above decomposition of L_0, take the sum and difference of the constraints (37) and reintroduce the Regge slope $\alpha' = \frac{1}{2}$ by dimensional analysis:

$$\begin{aligned} \alpha' M^2 &= 2(N + \tilde{N} - 2) \;, \\ N &= \tilde{N} \;. \end{aligned} \tag{41}$$

The first equation is the mass formula for string states, whereas the second equation shows that left- and right-moving degrees of freedom must contribute equally to the mass.

Let us list the lightest states satisfying these constraints:

Occupation	Mass	State
$N = \tilde{N} = 0$	$\alpha' M^2 = -4$	$\lvert k\rangle$
$N = \tilde{N} = 1$	$\alpha' M^2 = 0$	$\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1}\lvert k\rangle$
$N = \tilde{N} = 2$	$\alpha' M^2 = 4$	$\alpha^\mu_{-2}\tilde{\alpha}^\nu_{-2}\lvert k\rangle$
		$\alpha^\mu_{-2}\tilde{\alpha}^\nu_{-1}\tilde{\alpha}^\rho_{-1}\lvert k\rangle$
		$\alpha^\mu_{-1}\alpha^\nu_{-1}\tilde{\alpha}^\rho_{-2}\lvert k\rangle$
		$\alpha^\mu_{-1}\alpha^\nu_{-1}\tilde{\alpha}^\rho_{-1}\tilde{\alpha}^\sigma_{-1}\lvert k\rangle$

(42)

The most obvious and disturbing fact is that the ground state is a tachyon, i.e., a state of negative mass squared. Since the mass squared of a scalar corresponds to the curvature of the potential at the critical point, we seem to be expanding around a maximum rather then a minimum of the potential. This signals that the bosonic closed string quantized in flat Minkowski space is unstable. It is a very interesting question whether there is a minimum of this potential which provides a stable ground state. Since the tachyon acquires a vacuum expectation value in this minimum, this is referred to as tachyon condensation. But since we will be mostly interested in superstring theories, where tachyons are absent, we will simply ignore the fact that our toy model has a tachyon.

The first excited state is massless, and on top of it we find an infinite tower of states with increasing mass. Since the mass scale of string theory presumably is very large, we will focus on the massless states. So far we only imposed the constraints (37). The other constraints

$$L_m|\Phi\rangle = 0\;, \quad \tilde{L}_m|\Phi\rangle = 0\;, \quad m > 0\;, \tag{43}$$

impose conditions on the polarizations of physical states. For the tachyon one gets no condition, while for the first excited level the constraints with $m = 1$ are non-trivial. Forming a general linear combination of basic states,

$$\zeta_{\mu\nu}\alpha^{\mu}_{-1}\tilde{\alpha}^{\nu}_{-1}|k\rangle\;, \tag{44}$$

the constraints (43) imply

$$k^{\mu}\zeta_{\mu\nu} = 0 = k^{\nu}\zeta_{\mu\nu}\;. \tag{45}$$

Since $\zeta_{\mu\nu}$ is the polarization tensor, we see that only states of transverse polarization are physical. To obtain the particle content, we have to extract the irreducible representations of the D-dimensional Poincaré group contained in physical $\zeta_{\mu\nu}$. There are three such representations: the traceless symmetric part describes a graviton $G_{\mu\nu}$, the trace part corresponds to a scalar, the dilaton Φ, and the third representation is an antisymmetric tensor $B_{\mu\nu}$. In order to disentangle the trace part, one needs to introduce an auxiliary vector $\overline{k}$, with the properties:

$$\overline{k}\cdot\overline{k} = 0\;, \quad k\cdot\overline{k} = -1\;. \tag{46}$$

(k is the momentum vector.) The polarization tensors of the graviton, dilaton and antisymmmetric tensor are:

$$\begin{aligned}
\zeta^{G}_{\mu\nu} &= \zeta_{(\mu\nu)} - \frac{1}{D-2}\zeta^{\rho}_{\rho}(\eta_{\mu\nu} - k_{\mu}\overline{k}_{\nu} - k_{\nu}\overline{k}_{\mu})\;,\\
\zeta^{\Phi}_{\mu\nu} &= \frac{1}{D-2}\zeta^{\rho}_{\rho}(\eta_{\mu\nu} - k_{\mu}\overline{k}_{\nu} - k_{\nu}\overline{k}_{\mu})\;,\\
\zeta^{B}_{\mu\nu} &= \zeta_{[\mu\nu]}\;,
\end{aligned} \tag{47}$$

where $\zeta_{(\mu\nu)} = \frac{1}{2}(\zeta_{\mu\nu} + \zeta_{\nu\mu})$ and $\zeta_{[\mu\nu]} = \frac{1}{2}(\zeta_{\mu\nu} - \zeta_{\nu\mu})$ are the symmetric and antisymmetric parts of $\zeta_{\mu\nu}$. Note that the prefactor $1/(D-2)$ is needed in order that the trace part is physical. Using explicit choices for $k, \overline{k}$ one can check that $\zeta^G_{\mu\nu}$ is the polarization tensor of a plane wave and transforms as a traceless symmetric tensor under transverse rotations.

As we discussed above, physical states are only defined up to the addition of null states, $|\Phi\rangle \sim |\Phi\rangle + |\Psi\rangle$. In the case at hand adding null states corresponds to adding states of longitudinal polarization, according to:

$$\begin{aligned}
\zeta_{(\mu\nu)} &\sim \zeta_{(\mu\nu)} + k_\mu \zeta_\nu + \zeta_\mu k_\nu \\
\zeta_{[\mu\nu]} &\sim \zeta_{[\mu\nu]} + k_\mu \xi_\nu - \xi_\mu k_\nu \ .
\end{aligned} \tag{48}$$

ζ_μ and ξ_μ are arbitrary vectors orthogonal to the momentum k^μ. Adding null states can be understood as a residual gauge transformation parametrized by ζ_μ, ξ_μ. By taking Fourier transforms we see that these are the standard gauge invariances of a graviton and of an antisymmetric tensor, respectively:

$$\begin{aligned}
G_{\mu\nu} &\sim G_{\mu\nu} + \partial_\mu \Lambda_\nu + \partial_\nu \Lambda_\mu \ , \\
B_{\mu\nu} &\sim B_{\mu\nu} + \partial_\mu A_\nu - \partial_\nu A_\mu \ .
\end{aligned} \tag{49}$$

A graviton is defined by taking the gravitational action and expanding the metric around a flat background. The gauge transformations are then infinitesimal reparametrizations which, in a flat background, act according to (49) on the metric. Note that our gauge transformations Λ_μ, A_μ have a vanishing divergence, because the corresponding polarization vectors are orthogonal to the momentum. The reason is that the Virasoro constraints automatically impose a generalized Lorenz gauge.

Thus far our identification of the symmetric traceless part of the state (44) as a graviton is based on the fact that this state has the same kinematic properties as a graviton in Einstein gravity. We will see later, after analyzing string interactions, that this extends to the dynamical properties.

Finally it is interesting to compare the results of old covariant quantization to those obtained in light cone quantization. In light cone quantization unitarity is manifest, but the Lorentz algebra of the quantum theory has an anomaly which only cancels in the critical dimension $D = 26$. Moreover, the normal ordering constant must take the value $a = 1$. Independently, the same value of a is obtained when computing the Casimir energy of the ground state using ζ-function regularization. One virtue of light cone quantization is that one can write down immediately all the physical states. A basis is provided by all states which can be created using transverse oscillators,

$$\alpha^{i_1}_{-m_1} \cdots \tilde{\alpha}^{j_1}_{-n_1} \cdots |k\rangle \ , \tag{50}$$

where $i_1, \ldots, j_1, \ldots = 1, \ldots, D-2$. What remains is to group these states into representations of the D-dimensional Poincaré group. Massless states are classified by the little group $SO(D-2)$. Since all states manifestly are tensors with respect to this subgroup, one immediately sees that the massless states are a graviton (traceless symmetric tensor), dilaton (trace) and antisymmetric tensor. For massive states the little group is the full rotation subgroup $SO(D-1)$. Using Young tableaux it is straightforward to obtain these from the given representations of $SO(D-2)$.

Open Strings. Having treated the closed bosonic string in much detail, we now describe the results for open strings. One finds the same critical dimension, $D = 26$, and the same value of the normal ordering constant, $a = 1$. The constraints read:

$$(L_0 - 1)|\Phi\rangle = 0 \;, \quad L_m|\Phi\rangle = 0 \;, \quad m > 0 \;. \tag{51}$$

L_0 can be decomposed as $L_0 = \frac{1}{2}p^2 + N$, where N is the number operator. The L_0-constraint gives the mass formula:

$$\alpha' M^2 = N - 1 \;. \tag{52}$$

Therefore the lowest states are:

Occupation	Mass	State
$N = 0$	$\alpha' M^2 = -1$	$\lvert k\rangle$
$N = 1$	$\alpha' M^2 = 0$	$\zeta_\mu \alpha^\mu_{-1}\lvert k\rangle$
$N = 2$	$\alpha' M^2 = 1$	$\zeta_{\mu\nu}\alpha^\mu_{-1}\alpha^\nu_{-1}\lvert k\rangle$
		$\zeta_\mu \alpha^\mu_{-2}\lvert k\rangle$

(53)

The other constraints impose restrictions on the polarizations. Whereas the groundstate is a tachyonic scalar, the massless state has the kinematic properties of a gauge boson: its polarization must be transverse,

$$\zeta_\mu k^\mu = 0 \;, \tag{54}$$

and polarizations proportional to k^μ correspond to null states,

$$\zeta_\mu \sim \zeta_\mu + \alpha k_\mu \;. \tag{55}$$

This is the Fourier transform of a $U(1)$ gauge transformation,

$$A_\mu \sim A_\mu + \partial_\mu \chi \;. \tag{56}$$

Whereas massless closed string states mediate gravity, massless open string states mediate gauge interactions.

Chan-Paton Factors. Open string theory has a generalization which has non-abelian gauge interactions. One can assign additional degress of freedom to the ends of the string, namely charges ('Chan-Paton factors') which transform in the fundamental and anti-fundamental (complex conjugated) representation of the group $U(n)$. The massless states then take the form

$$\zeta_\mu \alpha^\mu_{-1}|k, a, \overline{b}\rangle \;, \tag{57}$$

where a is an index transforming in the fundamental representation $[n]$ of $U(n)$, whereas $\bar{b}$ transforms in the anti-fundamental representation $[\overline{n}]$. Since

$$[n] \times [\overline{n}] = \text{adj } U(n) , \tag{58}$$

the massless states transform in the adjoint of $U(n)$ and can be interpreted as $U(n)$ gauge bosons. (As for the graviton, we have only seen the required kinematic properties so far. But the interpretation is confirmed when studying interactions.)

Note that $U(n)$ is the only compact Lie group where the adjoint representation is the product of the fundamental and anti-fundamental representation. Therefore the construction precisely works for these groups.

Non-oriented Strings. There is a further modification which leads to non-oriented strings. These are obtained from the theories constructed so far by a projection. Both closed and open bosonic string theories are symmetric under world-sheet parity, which is defined as a reflection on the world-sheet:

$$\Omega : \sigma^1 \longrightarrow \pi - \sigma^1 = -\sigma^1 \;\text{ modulo }\; \pi . \tag{59}$$

Since Ω is an involution, $\Omega^2 = 1$, the spectrum can be organized into states with eigenvalues ± 1:

$$\Omega |N, k\rangle \quad = (-1)^N |N, k\rangle , \tag{60}$$

$$\Omega |N, \tilde{N}, k\rangle = |\tilde{N}, N, k\rangle . \tag{61}$$

Here $|N, k\rangle$ is an open string state with momentum k and total occupation number N and $|N, \tilde{N}, k\rangle$ is a closed string state with momentum k and total left and right occupation numbers $N, \tilde{N}$.

Non-oriented strings are defined by keeping only those states which are invariant under Ω. The resulting theories are insensitive to the orientation of the world-sheet. Let us look at the effect of this projection on the lowest states. For open strings we are left with:

Occupation	Mass	State
$N = 0$	$\alpha' M^2 = -1$	$\|k\rangle$
$N = 1$	$\alpha' M^2 = 0$	$-$
$N = 2$	$\alpha' M^2 = 1$	$\zeta_{\mu\nu} \alpha^{\mu}_{-1} \alpha^{\nu}_{-1} \|k\rangle$ $\zeta_{\mu} \alpha^{\mu}_{-2} \|k\rangle$

(62)

All states with odd occupation numbers are projected out, including the gauge boson. For closed strings we obtain:

Occupation	Mass	State
$N = \tilde{N} = 0$	$\alpha' M^2 = -4$	$\|k\rangle$
$N = \tilde{N} = 1$	$\alpha' M^2 = 0$	$\zeta_{(\mu\nu)} \alpha^{\mu}_{-1} \tilde{\alpha}^{\nu}_{-1} \|k\rangle$
$N = \tilde{N} = 2$	$\alpha' M^2 = 4$	$\zeta_{(\mu\nu)} \alpha^{\mu}_{-2} \tilde{\alpha}^{\nu}_{-2} \|k\rangle$
		$\zeta_{(\mu\rho\nu\sigma)} \alpha^{\mu}_{-1} \alpha^{\nu}_{-1} \tilde{\alpha}^{\rho}_{-1} \tilde{\alpha}^{\sigma}_{-1} \|k\rangle$

(63)

Only states which are left-right symmetric survive. At the massless level the antisymmetric tensor is projected out, whereas the graviton and dilaton are kept.

Chan-Paton Factors for Non-oriented Strings. The above construction can be generalized to open strings with Chan-Paton factors. In this case the two representations assigned to the ends of the strings must be equivalent. One can define a generalized involution Ω', which combines world-sheet parity with an action on the Chan-Paton indices,

$$\Omega' |N, a, b\rangle = \varepsilon (-1)^N |N, b, a\rangle \;, \tag{64}$$

where $\varepsilon = \pm 1$. The projection is $\Omega' |N, a, b\rangle \stackrel{!}{=} |N, a, b\rangle$. There are two inequivalent choices of the projection. For $\varepsilon = 1$, the indices a, b must transform in the fundamental representation of $SO(n)$. Since the adjoint of $SO(n)$ is the antisymmetric product of the fundamental representation with itself, the massless vector state transforms in the adjoint. More generally, states at even (odd) mass level transform as symmetric (antisymmetric) tensors.

The other choice is $\varepsilon = -1$. Then a, b transform in the fundamental of $USp(2n)$ (the compact form of the symplectic group). Our normalization is such that $USp(2) \simeq SU(2)$. Since the adjoint of $USp(2n)$ is the symmetric product of the fundamental representation with itself, the massless vector transforms in the adjoint. More generally, states at even (odd) mass level transform as antisymmetric (symmetric) tensors.

D-Branes. Finally we can consider open strings with Dirichlet boundary conditions along some directions. Consider first oriented open strings ending on a D-p-brane located at $x^a_{(1)}$. The ground state is tachyonic. The massless state of an open string with purely Neumann condition is a D-dimensional gauge boson $\alpha^{\mu}_{-1}|k\rangle$. Now we impose Dirichlet boundary conditions along the directions $a = p+1, \ldots, D-1$, so that the string can only move freely along the Neumann directions $m = 0, 1, \ldots, p$. The relevant kinematic group is now the world-volume Lorentz group $SO(1, \mathrm{p})$. The massless states are a world-volume vector,

$$\alpha^{m}_{-1}|k\rangle \;, \quad m = 0, 1, \ldots, \mathrm{p} \tag{65}$$

and $\mathrm{D}-\mathrm{p}-1$ scalars,

$$\alpha_{-1}^{a}|k\rangle\ , \quad a=\mathrm{p}+1,\ldots,\mathrm{D}-1\ . \tag{66}$$

The scalars correspond to transverse oscillations of the brane. Changing the position of the brane corresponds to changing the vacuum expectation values of the scalars. The effective action of the massless modes is given, to leading order in α', by the dimensional reduction of the D-dimensional Maxwell action to $\mathrm{p}+1$ dimensions. The full effective action is of Born-Infeld type.

Next consider N parallel D-p-branes, located at positions $x^{a}_{(i)}$. The new feature of this configuration is that there are strings which start and end on different branes. For such strings there is an additional term in the mass formula, which accounts for the stretching:

$$\alpha' M^2 = N - 1 + \left(\frac{|\boldsymbol{x}_{(i)}-\boldsymbol{x}_{(j)}|}{2\pi\sqrt{\alpha'}}\right)^2\ . \tag{67}$$

Here $\boldsymbol{x}_{(i)}$ is the position of the i-th brane. (Remember that the tension of the string is $(2\pi\alpha')^{-1}$.) Due to the normal ordering constant, the ground state becomes tachyonic if two branes come close enough. This signals an instability of the D-brane configuration. As already mentioned this might lead to interesting dynamics (tachyon condensation, decay of D-branes), but we will not discuss this here. Instead, we focus on features shared by D-branes in supersymmetric string theories. The states at the first excited level become massless precisely if the corresponding D-branes are put on top of each other. Each of the N branes already carries a $U(1)$ gauge theory: the massless modes of strings beginning and ending at the same brane give N vectors and $N\cdot(\mathrm{D}-\mathrm{p}-1)$ scalars. For N coinciding branes we get additional $N\cdot(N-1)$ vectors and $N\cdot(N-1)\cdot(\mathrm{D}-\mathrm{p}-1)$ scalars. Combining all massless states one gets one vector and $\mathrm{D}-\mathrm{p}-1$ scalars in the adjoint representation of the non-abelian group $U(N)$. This suggests that the D-brane system describes a $U(N)$ gauge theory with an adjoint Higgs mechanism. The Higgs mechanism is realized geometrically: Higgs expectation values correspond to the distances between branes, and the masses can be understood in terms of stretched strings. Again, this interpretation, which is based on analyzing the spectrum is confirmed when studying interactions. Besides Chan-Paton factors, D-branes are a second possibility to introduce non-abelian gauge groups. In fact Chan-Paton factors are related to D-branes through T-duality, but we will not be able to discuss this in these lectures.

The above construction can be extended to non-oriented strings, where other gauge groups occur. There are various other generalizations, which allow one to construct and study various gauge theories using strings and D-branes. These techniques are known as 'D-brane engineering' of field theories. Besides being of interest for the study of field theories through string methods, D-branes are important for understanding string theory itself. As we will see later, D-branes are actually solitons of string theory. Thus we are in the privileged position of knowing the exact excitation spectrum around such solitons in terms of open strings. This can be used, for example, to compute the entropy and Hawking radiation of black holes.

Another application of D-branes goes under the name of 'brane worlds' or 'brane universes' or 'models with large extra dimensions'. As we have seen, D-branes enable one to localize gauge interactions and matter on a lower-dimensional submanifold of space-time. This leads to models with space-dimensions where only gravity (closed strings) but not standard model matter (open strings) can propagate. Empirical limits on the size of the dimensions transverse to the brane only come from gravity, which is much weaker than all other interactions. Therefore such dimensions can be quite large, even up to about 1 mm. This is in contrast to limits on extra dimensions which are accessible to standard model interactions. Here the experimental limits are set by the scale resolvable in current accelerator experiments.

Brane world models are nowadays popular in both particle physics and cosmology. In particular, they can be used to construct models where the fundamental gravitational scale is of order 1 TeV. We will come back to these applications of D-branes in Sect. 6.

2.3 Further Reading

The material covered in this section can be found in all of the standard textbooks on string theory [1–5]. Dirichlet boundary conditions and D-branes are only covered by the more recent ones [3,4].

3 Interacting Bosonic Strings

So far we have not specified how strings interact. One might expect that this can be done by adding interaction terms to the world-sheet action. However, we have to respect the local symmetries of the Polyakov action, which severely restricts our options. In particular, contact interactions, which are frequently used in describing non-fundamental string-like objects such as polymers, are not compatible with Weyl invariance. Admissible interacting world-sheet actions include marginal deformations of the Polyakov action, i.e., deformations which preserve Weyl invariance. One such deformation replaces the flat space-time metric by a curved one. As expected intuitively, such an action does not describe interactions among strings, but strings moving in a non-trivial background. The same is true when replacing the Polaykov action by more general conformal field theories.

How then do we define interactions? We will give a heuristic discussion in the next section. The resulting scattering amplitudes are Lorentz covariant, unitary and UV finite. They include the Veneziano amplitude and its cousins, which historically started the subject.

For definiteness we will focus in the following on closed oriented strings. The generalization to other string theories will be indicated briefly.

3.1 Heuristic Discussion

Intuitively, interactions between strings are described by world-sheets which connect a given initial configuration of strings to a final configuration. One can draw several such world-sheets, which differ by their topologies. Comparing to the similar treatment of point particles by graphs, we realize that while graphs have vertices, the world-sheets connecting strings are manifolds without distinguished interaction points. This leads to the expectation that string interactions are less singular then those of point particles, which is indeed confirmed by the final result of the construction. Moreover, it indicates that one does not have any freedom in defining interactions. For particles, we can assign couplings to vertices which depend on the species of the particles meeting at the vertex. For strings the interaction is encoded in the topology of the world-sheet and there is no such freedom. There is one fundamental interaction, which couples three closed strings, and all we can do is to assign a coupling constant κ to it.

Next, we restrict ourselves to finding transition amplitudes between asymptotic states in the infinite past and future. An asymptotic in- or out-going state is represented by a semi-infinite cylinder. When mapping this to a punctered disc, the asymptotic state is represented by the puncture. This leads to the idea that we can represent the asymptotic state by a local operator of the world-sheet field theory. Such operators are called vertex operators. Note that they do not describe interactions. Instead, the vertex operator $V_{\Phi}(\sigma)$ describes the creation or annihilation of the string state $|\Phi\rangle$ at the position σ on the world-sheet. That is, they allow us to assign a copy of the space of physical states to every point of the world-sheet. As we will see below, there is indeed a natural one-to-one map between physical states $|\Phi\rangle$ and local operators of the world-sheet field theory.

After replacing the world-sheet punctures by insertions of vertex operators we are left with compact closed surfaces. The topologies of such surfaces are classified by their genus $g \geq 0$, or equivalently, by their Euler number $\chi = 2-2g$. Here $g=0$ is the two-sphere, and $g=1$ is the torus. The general genus g surface Σ_g is obtained from the sphere by attaching g handles. The handles play the role of loops in Feynman diagrams. When considering an interaction process on Σ_g involving M external states, we find $M-\chi$ fundamental string interactions and have to assign a factor $\kappa^{M-\chi}$.

We now postulate that a scattering amplitude involving M external states is given by

$$A(1,\dots,M) = \sum_{g=0}^{\infty} \kappa^{M-\chi} A(1,\dots,M)_g \ , \tag{68}$$

where $A(1,\dots,M)_g$ is the contribution of Σ_g. This is a perturbative expression in the string coupling κ. As usual for theories with a single coupling, the expansion in the coupling coincides with the expansion in loops, which in our case is the expansion in the genus g.

The genus g contribution is defined to be

$$A(1,\dots,M)_g = \langle V_1 \cdots V_M \rangle_g \ , \tag{69}$$

where

$$V_i = \int_{\Sigma_g} d^2\sigma_i \sqrt{h} V_i(\sigma_i) \tag{70}$$

are the so-called integrated vertex operators, which are obtained by integrating the vertex operators $V_i(\sigma_i)$ over the world sheet. (Though our notation might suggest otherwise, we do not require that Σ_g can be covered by one set of coordinates, which is of course impossible for compact Σ_g. We just use a local representative of the integrand for notational purposes.) In (69) we compute the correlation function of the vertex operators $V_i(\sigma_i)$ on Σ_g in the world-sheet quantum field theory defined by the Polyakov action and integrate over the positions of the vertex operators. The result is interpreted as a scattering amplitude of string states in space-time, with the in- and out-states represented by the vertex operators.

Note that it is not possible to introduce arbitrary weight factors between the contributions of different genera. The reason is that unitarity requires that scattering amplitudes factorize into the amplitudes of subprocesses whenever an intermediate state is on-shell. In fact, in the old days of string theory this was used to construct the perturbative expansion by seewing together tree amplitudes. However, this approach is more cumbersome then the Polyakov path integral approach that we will describe here.

3.2 Vertex Operators

We now take a closer look at the vertex operators. Observe that the scattering amplitudes defined by (68,69,70) must be invariant under reparametrizations of the world-sheets. In particular the local vertex operators $V_i(\sigma_i)$ must transform such that (70) is invariant. When imposing the conformal gauge, it still must transform in a specific way under conformal transformations $\sigma^\pm \rightarrow \tilde{\sigma}^\pm(\sigma^\pm)$. In conformal field theory fields which transform covariantly under conformal transformations are called primary conformal fields. A primary conformal field of weights $(h, \overline{h})$ is an object that transforms like a contravariant tensor field of rank $(h, \overline{h})$:

$$\tilde{V}(\tilde{\sigma}^+, \tilde{\sigma}^-) = \left(\frac{d\sigma^+}{d\tilde{\sigma}^+} \right)^h \left(\frac{d\sigma^-}{d\tilde{\sigma}^-} \right)^{\overline{h}} V(\sigma^+, \sigma^-) \,. \tag{71}$$

Invariance of (70) implies that vertex operators of physical states must be primary conformal fields of weights $(1, 1)$. This property is equivalent to imposing the Virasoro constraints (41) on physical states. States assigned to a point P of Σ are constructed from vertex operators by applying them to a ground state $|0\rangle_P$,

$$|\Phi\rangle = V_\Phi(P)|0\rangle_P \,. \tag{72}$$

To make contact with the space $\mathcal{F}_{\text{phys}}$ constructed in Sect. 2.2, one parametrizes Σ in the vicinity of P by a semi-infinity cylinder with P being the asymptotic point $\sigma^0 \to -\infty$. Intuitively this describes an ingoing state created in the infinite past. Then,

$$|\Phi\rangle = \lim_{\sigma^0 \to -\infty} V_\Phi(\sigma)|0\rangle \ , \tag{73}$$

where $|0\rangle := |k = 0\rangle$ is the (unphysical) zero-momentum state with occupation numbers $N = 0 = \tilde{N}$ in $\mathcal{F}$.

To indicate how this works in practice, we now specify the vertex operators for the lowest states. Consider the operator

$$V(\sigma) =: e^{\mathrm{i}k_\mu X^\mu} : (\sigma) \ , \tag{74}$$

where $: \cdots :$ indicates normal ordering. Applying this operator we find

$$\lim_{\sigma^0 \to -\infty} : e^{\mathrm{i}k_\mu X^\mu} : (\sigma)|0\rangle = e^{\mathrm{i}k_\mu x^\mu}|0\rangle = |k\rangle \ , \tag{75}$$

where we have used that $e^{\mathrm{i}k_\mu x^\mu}|0\rangle$ is an eigenstate of p^μ with eigenvalues k^μ. One can show that (74) has weights $(\frac{1}{8}k^2, \frac{1}{8}k^2)$. Thus it has weights $(1,1)$ if $k^2 = 8$, which is the physical state condition $M^2 = -8$ for the tachyonic ground state of the closed string. (We have set $\alpha' = \frac{1}{2}$.)

The vertex operator for the first excited level is

$$V(\sigma) =: \zeta_{\mu\nu}\partial_+ X^\mu \partial_- X^\nu e^{\mathrm{i}k_\rho X^\rho} : (\sigma) \ . \tag{76}$$

This has weights $(1,1)$ if

$$k^2 = 0 \ , \quad k^\mu \zeta_{\mu\nu} = 0 = k^\nu \zeta_{\mu\nu} \ , \tag{77}$$

which is precisely the physical state condition for the state

$$\zeta_{\mu\nu}\alpha^\mu_{-1}\tilde{\alpha}^\nu_{-1}|k\rangle \ . \tag{78}$$

More generally, vertex operators of the form

$$V(\sigma) =: \zeta_{\mu_1\cdots\nu_1\cdots}\partial_+^{m_1} X^{\mu_1} \cdots \partial_-^{n_1} X^{\nu_1} \cdots e^{\mathrm{i}k_\rho X^\rho} : (\sigma) \tag{79}$$

generate states of the form

$$\zeta_{\mu_1\cdots\nu_1\cdots}\alpha^{\mu_1}_{-m_1}\cdots\tilde{\alpha}^{\nu_1}_{-n_1}\cdots|k\rangle \ . \tag{80}$$

3.3 Interactions in the Path Integral Formalism

The next step is to explain in more detail how the amplitudes (68)–(70) are defined and how they are computed in practice. As usual one can use either the path integral (Lagrangian) or the operator (Hamiltonian) formulation. We will use Polyakov's path integral formulation. This has the advantage of immediately providing explicit formal expressions for correlation functions. The mathematical complications of defining the interacting quantum theory are hidden in the path integral measure. We will not discuss this in full detail, but mention and illustrate the most important points.

The Path Integral. We now turn to the Polyakov path integral, which is one way to give a precise meaning to (68). In this approach the correlation function (69) is computed by functional methods. Intuitively we integrate over all paths that strings can take in space-time. However, in order to have a well defined path integral, we need to study the theory in Euclidean signature, both on the world-sheet and in space-time. A Euclidean formulation of the world-sheet theory is needed to have a well defined functional integral for the world-sheet field theory. In particular, we want to have well defined world-sheet metrics on general surfaces Σ_g, which is not possible for Lorentzian signature. Second, one also has to work in Euclidean space-time, in order to have a standard Gaussian integral for the 'time' coordinate X^0. Wick-rotating X^0 can be interpreted as continuing to unphysical Euclidean momenta and polarizations. As we have seen in our discussion of vertex operators the string coordinates X^μ are always contracted with momenta and polarizations. Physical scattering amplitudes are thus obtained by computing (68) in the Euclidean theory and evaluating the result for physical momenta and polarizations. This uses the analycity properties expected to hold for any relativistic unitary scattering amplitude. For tree-level amplitudes one can study how the Wick rotation works explicitly, by comparing to results obtained by operator methods.

Our starting point is the Polyakov action on a world-sheet Σ with positive definite metric $h_{\alpha\beta}$ and local complex coordinate z,

$$S_P = \frac{1}{4\pi\alpha'} \int_\Sigma d^2 z \sqrt{h} h^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X_\mu \ . \tag{81}$$

The quantum theory is now defined by summing over all topologies of Σ and integrating over X^μ and $h_{\alpha\beta}$:

$$A(1, \ldots , M) = \sum_{g=0}^{\infty} \kappa^{M-\chi} N_g \int DX^\mu D h_{\alpha\beta} e^{-S_P[X,h]} V_1 \cdots V_M \ , \tag{82}$$

where V_i are the integrated vertex operators of the physical states and N_g are normalization factors needed to define the path integral. The V_i depend on X^μ through the local vertex operators $V_i(\sigma_i)$, while the world-sheet metric enters through the integration over σ_i.

One expects that one can properly define and compute the expression (82), because the integration over X^μ is Gaussian (in flat space-time) and $h_{\alpha\beta}$ is non-dynamical. This turns out to be true, though several interesting complications arise. Let us consider the integration over $h_{\alpha\beta}$. Since we can locally impose the conformal gauge,

$$h_{\alpha\beta} = \delta_{\alpha\beta} \ , \tag{83}$$

we expect that we can use the Faddeev-Popov method and trade the integration over the metric for an integration over reparametrizations and the Weyl factor. As long as these are symmetries, the corresponding integration factorizes and can be absorbed in the normalization factor N_g. The first obstruction encountered is that there is a conformal anomaly when the quantum theory based on (81) lives on a curved world-sheet. This has the consequence that the integration over the Weyl factor does not factorize in general. One option is to accept it as a new, purely quantum degree of freedom: this is non-critical string theory, also called Liouville string theory, because the dynamics of the Weyl factor is given by the Liouville action. The other option is to observe that the anomaly is proportional to $D-26$, and therefore cancels for $D=26$ space-time dimensions. This is the critical string theory we study in these lectures.

Moduli and Modular Transformations. The next point is that the gauge (83) cannot be imposed globally. All that can be achieved is to map $h_{\alpha\beta}$ to a metric of constant curvature,

$$h_{\alpha\beta} \stackrel{!}{=} \hat{h}_{\alpha\beta}[\boldsymbol{\tau}] \; . \tag{84}$$

As indicated, Σ_g in general possesses a continuous family of such metrics, parametrized by moduli $\boldsymbol{\tau} = (\tau_1, \ldots)$. The space of constant curvature metrics on a two-dimensional closed compact surface is isomorphic to the space of complex structures. By reparametrizations and Weyl transformations we cannot change the complex structure of the metric but we can map it to the unique representative (84) of the complex structure class which has constant curvature. Then the path integral over all metrics reduces to a finite-dimensional integral over the space $\mathcal{M}_g$ of complex structures. The dimension of this space is known from the Riemann-Roch theorem. For $g=0$ the complex structure is unique, and every metric can be mapped to the standard round metric on the sphere. For $g>1$ there is a non-trivial moduli space,

$$\begin{aligned} \dim_{\mathbf{C}} \mathcal{M}_g &= 1 \; , \quad \text{for} \;\; g=1 \; , \\ \dim_{\mathbf{C}} \mathcal{M}_g &= 3g-3 \; , \quad \text{for} \;\; g>1 \; . \end{aligned} \tag{85}$$

After carrying out the integration over the metric, amplitudes take the form

$$A(1,\ldots,M) = \sum_{g=0}^{\infty} \kappa^{M-\chi} N_g' \int_{\mathcal{M}_g} d\mu(\boldsymbol{\tau}) \int DX^\mu e^{-S_P[X,\hat{h}]} J(\hat{h}) V_1 \ldots V_M \; . \tag{86}$$

N_g' are normalization factors needed to deal with the X^μ-integration and $J(\hat{h})$ is the Faddeev-Popov determinant, which one can rewrite as a functional integral over Faddeev-Popov ghost fields. As indicated the X^μ-integral depends on the moduli through the world-sheet metric $\hat{h}_{\alpha\beta} = \hat{h}_{\alpha\beta}(\boldsymbol{\tau})$. One finds that the measure $d\mu(\boldsymbol{\tau})$ for the moduli is the natural measure on the space of complex structures, the so-called Weil-Petersson measure.

The precise characterization of the moduli space has further interesting details. We examplify this with the two-torus. We can represent a torus as a parallelogram in the complex plane with opposite sides identified. Since the complex structure does not depend on the overall volume, we can restrict ourselves to parallelograms with edges $0, 1, \tau, \tau+1$, where $\mathrm{Im}(\tau) > 0$. In one complex dimension holomorphic maps are conformal maps, and vice versa. Thus the complex structure is varied by moving τ in the upper half-plane,

$$\mathcal{H} = \{\tau \in \mathbf{C} \mid \mathrm{Im}(\tau) > 0\} \ . \tag{87}$$

This is the modulus we are looking for. $\mathcal{H}$ has a metric of constant negative curvature, the Poincaré metric,

$$d\mu(\tau) = \frac{d^2\tau}{(\mathrm{Im}(\tau))^2} \ . \tag{88}$$

With this $Sl(2,\mathbf{R})$–invariant metric, $\mathcal{H}$ is the symmetric space $Sl(2,\mathbf{R})/SO(2)$. However, $\mathcal{H}$ is not our moduli space, because it overcounts complex structures. On $\mathcal{H}$ the group $Sl(2,\mathbf{R})$ acts from the right. Taking τ as coordinate, the operation is

$$\tau \longrightarrow \frac{a\tau+b}{c\tau+d} \ , \quad \text{where} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Sl(2,\mathbf{R}) \ . \tag{89}$$

The subgroup $Sl(2,\mathbf{Z})$ maps parallelograms to parallelograms which define the same torus, because they form basic cells of the same lattice in $\mathcal{H}$. Such transformations are called modular transformations. Their action on the torus is given by cutting the torus along a non-contractible loop, twisting and regluing. This corresponds to a large reparametrization which cannot be continously connected to the identity. Clearly, we have to require that string amplitudes are invariant under such large reparametrizations. This implies a consistency condition, known as modular invariance: the $\boldsymbol{\tau}$-integral in (86) must be invariant under modular transformations. This condition becomes non-trivial when considering more general background geometries or string theories with fermions.

The moduli space is obtained by restricting to a fundamental domain $\mathcal{F}$ of the action of $Sl(2,\mathbf{Z})$ on $\mathcal{H}$. By modular invariance we can consistently restrict the τ-integration to such an $\mathcal{F}$. The standard choice is found by looking at the action of the two generators of $Sl(2,\mathbf{Z})$,

$$\tau \to \tau + 1 \ , \quad \tau \to -\frac{1}{\tau} \ . \tag{90}$$

Therefore the most convenient choice is

$$\mathcal{F} = \{\tau \in \mathcal{H} | -\tfrac{1}{2} \leq \mathrm{Im}(\tau) < \tfrac{1}{2}\tau \ , \quad |\tau| \geq 1\} \tag{91}$$

(with certain identifications along the boundary).

Modular invariance has deep consequences for the short distance behaviour of string theory. In fact, modular invariance is what makes closed string theories UV finite. To illustrate how this works, note that a one-loop amplitude in closed string theory takes the form

$$A^{\text{String}}_{1-\text{loop}} \sim \int_{\mathcal{F}} \frac{d^2\tau}{(\mathrm{Im}(\tau))^2} F(\tau) \ . \tag{92}$$

An analogous expression for one loop amplitudes in quantum field theory is given by Schwinger's proper time parametrization,

$$A^{\mathrm{QFT}}_{1-\mathrm{loop}} \sim \int_{\varepsilon}^{\infty} \frac{dt}{t} f(t) , \tag{93}$$

where t is the proper time and ε is an UV cutoff. In this formulation UV divergencies occur at short times $t \to 0$. In string theory $\mathrm{Im}(\tau)$ plays the role of proper time, and potential UV divergencies occur for $\mathrm{Im}(\tau) \to 0$. However, by restricting to the fundamental domain we have cut out the whole dangerous region of small times and high momenta. This confirms the intuitive idea that strings should have a particularly soft UV behaviour, because the theory has a minimal length scale, which works like a physical UV cutoff. Note that one still has IR divergencies. In bosonic string theory one has divergencies related to the tachyon, which show that the theory is unstable in Minkowski space. This problem is absent in supersymmetric string theories. In addition one can have IR divergencies related to massless states. Since there is only a finite number of massless string states, this problem has the same character as in field theory.

Also note that the modular transformation $\tau \to -1/\tau$ maps the UV region of $\mathcal{H}$ to its IR region. Thus, modular transformations map UV divergencies to IR divergencies and enable us to reinterpret them in terms of low energy physics (namely, intermediate massless states which go on-shell).

For higher genus surfaces Σ_g with $g > 1$ the story is similar, but more complicated. There is an analogue of the upper half plane, which is called Siegels upper half plane and has complex dimension $\frac{g(g+1)}{2}$. Since there are only $3g-3$ complex moduli, this space contains more parameters then needed for $g \geq 4$. The Teichmüller space is embedded in a complicated way into Siegel's upper half plane. On top of this there is a modular group which has to be divided out.

Global Conformal Transformations. The integration over complex structure moduli in (86) reflects that surfaces with $g > 0$ have metrics that cannot be related by reparametrizations. Therefore there is a finite left-over integration when replacing the integral over metrics by an integral over reparametrizations. For $g < 2$ one has in addition the reciprocal phenomenon: these surfaces have global conformal isometries. This means that there are reparametrizations which do not change the metric, implying an overcounting of equivalent contributions in (86). Formally this is taken care of by the normalization factors N_0', N_1'. The overcounting yields a multiplicative factor, which is the volume of the group of conformal isometries. This has to be cancelled by the normalization factors. For $g = 0$ the conformal group is $Sl(2, \mathbf{C})$ and has infinite volume. Thus one has to formally divide out an infinite constant. For $g = 1$ the conformal group is $U(1)^2$, and has a volume which depends on the complex structure modulus τ of the world-sheet. This factor is crucial for world-sheet modular invariance.

The systematic approach is to treat the global conformal isometries as a residual gauge invariance and to apply the Faddeev-Popov technique.Then the volumes of residual gauge groups are properly taken care of. So far we have been sloppy about how and when to carry out the integration over the positions of the vertex operators. The proper order is as follows: one first carries out the X^μ-integration to obtain a correlation function on a world-sheet of given topology and complex structure:

$$\langle V_1(z_1, \overline{z}_1) \cdots \rangle_{g,\tau} = N'_g \int DX e^{-S_P[X, \hat{h}(\boldsymbol{\tau})]} J(\hat{h}(\boldsymbol{\tau})) V_1(z_1, \overline{z}_1) \cdots . \tag{94}$$

Next one integrates over the positions of the vertex operators. For $g < 2$ one treats the global conformal isometries by the Faddeev Popov method. The result is

$$\langle V_1 \cdots \rangle_{g,\boldsymbol{\tau}} = \int d\mu(z_1, \overline{z}_1, \dots) \langle V_1(z_1, \overline{z}_1) \cdots \rangle_{g,\boldsymbol{\tau}} , \tag{95}$$

where $d\mu(z_1, \overline{z}_1, \dots)$ for $g < 2$ is a measure invariant under the global isometries.

For $g = 0$ the measure vanishes if less than three vertex operators are present. This reflects the infinite volume of the global conformal group: by $Sl(2, \mathbf{C})$ transformations one can map three points on the sphere to three arbitrary prescribed points. Thus, the $Sl(2, \mathbf{C})$ symmetry can be used to keep three vertex operators at fixed positions. In other words the first three integrations over vertex operators compensate the infinite volume of the global conformal group that one has to divide out. For less than three vertex operators one cannot compensate this infinite normalization factor and the result is zero. Thus, the integrated zero-, one- and two-point functions vanish. This implies that at string tree level the cosmological constant and all tadpoles diagrams vanish.

The final step in evaluating (86) is to integrate over complex structures and to sum over topologies:

$$A(1, \dots, M) = \sum_{g=0}^{\infty} \kappa^{M - \chi(g)} \int_{\mathcal{M}_g} d\mu(\boldsymbol{\tau}) \langle V_1 \dots \rangle_{g,\boldsymbol{\tau}} . \tag{96}$$

Through the vertex operators, $A(1, \dots, M)$ is a function of the momenta k_i^μ and polarization tensors $\zeta_i^{\mu_1 \cdots}$ of the external states.

Graviton Scattering. Though we cannot go through the details of a calculation here, we would like to discuss the properties of string scattering amplitudes in a particular example. Our main interest being gravity, we choose the scattering of two massless closed string states. The corresponding external states are

$$\zeta^{(i)}_{\mu\nu} \alpha^\mu_{-1} \tilde{\alpha}^\nu_{-1} | k^{(i)} \rangle , \tag{97}$$

with $i = 1, 2, 3, 4$. The resulting amplitude takes the following form:

$$A_4^{\text{String}} = \kappa^2 \frac{\Gamma(-\frac{\alpha'}{4}s)\Gamma(-\frac{\alpha'}{4}t)\Gamma(-\frac{\alpha'}{4}u)}{\Gamma(1+\frac{\alpha'}{4}s)\Gamma(1+\frac{\alpha'}{4}t)\Gamma(1+\frac{\alpha'}{4}u)} \cdot K(\zeta^{(i)}, k^{(i)}) . \tag{98}$$

Here s, t, u are the Mandelstam variables

$$s = -(k^{(1)} + k^{(2)})^2 \ , \quad t = -(k^{(2)} + k^{(3)})^2 \ , \quad u = -(k^{(1)} + k^{(3)})^2 \tag{99}$$

and $K(\zeta^{(i)}, k^{(i)})$ is the kinematic factor, a complicated function of momenta and polarizations that we do not display.

Scattering amplitudes have poles whenever an intermediate states can be produced as a real physical state. Unitarity requires that the residue of the pole describing such a resonance is the product of the amplitudes of the subprocesses through which the intermediate state is produced and decays. In this way the pole structure of amplitudes is related to the particle spectrum of the theory.

The amplitude (98) has poles when the argument of one of the Γ-functions in the numerator takes a non-positive integer value,

$$-\frac{\alpha'}{4} x = 0, -1, -2, \dots \ , \quad \text{where } \ x = s, t, u \ . \tag{100}$$

Comparing to the mass formula of closed strings,

$$\alpha' M^2 = 2(N + \tilde{N} - 2) \ , \tag{101}$$

we see that the poles precisely correspond to massless and massive string states with $N = \tilde{N} = 1, 2, 3, \dots$. There is no pole corresponding to the tachyon ($N = 0$) in this amplitude, because the tachyon cannot be produced as a resonance for kinematic reasons. When computing the amplitude for tachyon scattering instead, one also finds a tachyon pole.

The particular pole structure of (98) and of related string amplitudes was observed before the interpretation of the amplitudes in terms of strings was known. In the late 1960s it was observed experimentally that hadronic resonances obey a linear relation between the spin and the square of the mass, called Regge behaviour. This behaviour was correctly captured by the Veneziano amplitude, which has a structure similar to (98) and describes the scattering of two open string tachyons. The Regge behaviour was the clue for the interpretation of the Veneziano amplitudes and its cousins in terms of strings.

To see that string states show Regge behaviour, consider the truncation of string theory to four space-time dimension (which is consistent at tree level). Closed string states with level $N = \tilde{N}$ have spins $J \leq 2N$, because the spin J representation of the four-dimensional Lorentz group is the traceless symmetric tensor of rank J. Open string states have spins $J \leq N$. The states lie on lines in the (M^2, J)–plane, which are called Regge trajectories. The closed string Regge trajectories are given by

$$\alpha_{\text{closed}}(M^2) = \alpha'_{\text{closed}} M^2 + \alpha_{\text{closed}}(0) \ , \tag{102}$$

where

$$\alpha'_{\text{closed}} = \frac{\alpha'}{4} \ , \quad \alpha_{\text{closed}}(0) = 1, 0, -1, \dots \ . \tag{103}$$

String states correspond to those points on the Regge trajectories where $\alpha'_{\text{closed}}(M^2) = N + \tilde{N}$. States with the maximal possible spin $J = 2N = N + \tilde{N}$ for a given mass lie on the leading Regge trajectory $\alpha_{\text{closed}}(0) = 1$. Since α' determines the slope of the trajectories, it is called the Regge slope. The corresponding expressions for open strings are:

$$\alpha_{\text{open}}(M^2) = \alpha'_{\text{open}} M^2 + \alpha_{\text{open}}(0) \;, \tag{104}$$

where

$$\alpha'_{\text{open}} = \alpha' \;, \quad \alpha_{\text{open}}(0) = 1, 0, -1, \ldots \;. \tag{105}$$

The resonances found in open string scattering lie on the corresponding Regge trajectories.

When computing scattering amplitudes in terms of Feynman diagrams in field theory, individual diagrams only have poles in one particular kinematic channel, i.e., in the s-channel or t-channel or u-channel. The full scattering amplitude, which has poles in all channels, is obtained by summing up all Feynman diagrams. In (closed oriented) string theory there is only one diagram in each order of perturbation theory, which simultanously has poles in all channels. The total amplitude can be written as a sum over resonances in one particular channel, say the s-channel. This is consistent with the existence of poles in the other channels, because there is an infinite set of resonances. When instead writing the amplitude in the form (98), it is manifestly symmetric under permutations of the kinematic variables s, t, u. This special property was called 'duality' in the old days of string theory (a term that nowadays is used for a variety of other, unrelated phenomena as well).

Another important property of (98) and other string amplitudes is that they fall off exponentially for large s, which means that the behaviour for large external momenta is much softer than in any field theory. This is again due to the presence of an infinite tower of excitations. Since loop amplitudes can be constructed by sewing tree amplitudes, this implies that the UV behaviour of loop diagrams is much softer than in field theory. This lead to the expectation that string loop amplitudes are UV finite, which was confirmed in the subsequent development of string perturbation theory.

Though we did not explicitly display the kinematic factor $K(\zeta^{(i)}, k^{(i)})$ we need to emphasize one of its properties: it vanishes whenever one of the external states is a null state. As we learned above, null states have polarizations of the form

$$\zeta^{(i)}_{\mu\nu} = k^{(i)}_\mu \xi^{(i)}_\nu + k^{(i)}_\nu \zeta^{(i)}_\mu \tag{106}$$

and are gauge degrees of freedom. They have to decouple from physical scattering amplitudes, as it happens in the above example. This property is called 'on shell gauge invariance', because it is the manifestation of local gauge invariance at the level of scattering amplitudes. It can be proven to hold for general scattering amplitudes.

If we take the polarization tensors of the external states to be symmetric and traceless, then (98) describes graviton–graviton scattering. So far our identification of this string state with the graviton was based on its kinematic properties. Since Einstein gravity is the only known consistent interaction for a second rank, traceless symmetric tensor field (‘massless spin–2–field’), we expect that this also holds dynamically. We will now check this explicitly. In the field theoretical perturbative approach to quantum gravity one starts from the Einstein-Hilbert action,

$$S = \frac{1}{2\kappa^2} \int d^D x \sqrt{g} R \tag{107}$$

and expands the metric around flat space

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \kappa \psi_{\mu\nu}(x) \ . \tag{108}$$

The field $\psi_{\mu\nu}(x)$ is the graviton field. Expanding (107) in κ one gets a complicated non-polynomial action for ψ that one quantizes perturbatively. The resulting theory is non-renormalizable, but tree diagrams can be consistently defined and computed. In particular one can compute graviton–graviton scattering at tree level and compare it to the string amplitude (98). Denoting the field theory amplitude by A_4^{FTh}, the relation is

$$A_4^{\mathrm{String}} = \frac{\Gamma(1 - \frac{\alpha'}{4}s)\Gamma(1 - \frac{\alpha'}{4}t)\Gamma(1 - \frac{\alpha'}{4}u)}{\Gamma(1 + \frac{\alpha'}{4}s)\Gamma(1 + \frac{\alpha'}{4}t)\Gamma(1 + \frac{\alpha'}{4}u)} A_4^{\mathrm{FTh}} \ . \tag{109}$$

In the limit $\alpha' \to 0$, which corresponds to sending the string mass scale to infinity, the string amplitude reduces to the field theory amplitude:

$$\lim_{\alpha' \to 0} A_4^{\mathrm{String}} = A_4^{\mathrm{FTh}} \ . \tag{110}$$

At finite α' string theory deviates from field theory. The correction factor in (109) contains precisely all the poles corresponding to massive string states, whereas the massless poles are captured by the field theory amplitude. One can construct an effective action which reproduces the string amplitude order by order in α'. At order α' one obtains four-derivative terms, in particular terms quadratic in the curvature tensor,

$$S_{\mathrm{eff}} = \frac{1}{2\kappa^2} \int d^D x \sqrt{g} (R + \alpha' c_1 R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} + \cdots + \mathcal{O}((\alpha')^2)) \ , \tag{111}$$

where c_1 is a numerical constant. The α'-expansion of the effective action is an expansion in derivatives. It is valid at low energies, i.e., at energies lower than the scale set by α', where corrections due to massive string scales are small.

Obviously, it is very cumbersome to construct the effective action by matching field theory amplitudes with string amplitudes. In practice one uses symmetries to constrain the form of the effective action. This is particularly efficient for supersymmetric actions, which only depend on a few independent parameters or functions, which can be extracted from a small number of string amplitudes. A different technique, which often is even more efficient, is to study strings in curved backgrounds, and, more generally, in background fields.

3.4 Strings in Curved Backgrounds

So far we only discussed strings in flat backgrounds. Let us now consider the case of a curved background with Riemannian metric $G_{\mu\nu}(X)$. Then the Polyakov action takes the form of a non-linear sigma-model

$$S_P = \frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{h} h^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X^\nu G_{\mu\nu}(X) \ . \tag{112}$$

As emphasized above, the local Weyl invariance

$$h_{\alpha\beta} \to e^{\Lambda(\sigma)} h_{\alpha\beta} \tag{113}$$

is crucial for the consistency of string theory, since the construction of states, vertex operators and amplitudes is based on having a conformal field theory on the world-sheet. If the space-time metric is curved, then the Weyl invariance of the classical action (112) is still manifest. But at the quantum level it becomes non-trivial and imposes restrictions on $G_{\mu\nu}(X)$. In the non-linear sigma-model defined by (112) one can define a modified beta function $\overline{\beta}$, which measures the violation of local Weyl invariance. In order to have local Weyl invariance this function must vanish,

$$\overline{\beta} = 0 \ . \tag{114}$$

Since $G_{\mu\nu}(X)$ are the field-dependent couplings of the non-linear sigma-model, the beta function $\overline{\beta}$ is a functional of $G_{\mu\nu}(X)$. It can be computed perturbatively, order by order in α'. The dimensionless expansion parameter is the curvature scale of the target space (i.e., space-time) measured in units of the string length $\sqrt{\alpha'}$.

The leading term in this expansion is:

$$\overline{\beta}^G_{\mu\nu} = -\frac{1}{2\pi} R_{\mu\nu} \ . \tag{115}$$

Thus the space-time background has to be Ricci-flat, i.e., it satisfies the vacuum Einstein equation. The condition imposed on the background field by local Weyl invariance on the world-sheet is its space-time equation of motion. This relation between world-sheet and space-time properties holds for other background fields as well and can be used as an efficient method to construct effective actions. One can also compute the α'-corrections to (115):

$$\overline{\beta}^G_{\mu\nu} = -\frac{1}{2\pi} \left(R_{\mu\nu} + \frac{\alpha'}{2} R_{\mu\alpha\beta\gamma} R_\nu{}^{\alpha\beta\gamma} \right) \ . \tag{116}$$

The corresponding α'-corrections to the Einstein-Hilbert action take the form (111).

At this point we need to reflect a little bit on how gravity is described in string theory. So far we have seen that it enters in two ways: first, there is a graviton state $\zeta_{(\mu\nu)}\alpha^{\mu}_{-1}\tilde{\alpha}^{\nu}_{-1}|k\rangle$ in the string spectrum. Second, there is a background metric $G_{\mu\nu}(X)$. If gravity is described consistently, then these two objects must be related. To explore this we expand $G_{\mu\nu}(X)$ around flat space,

$$G_{\mu\nu}(X) = \eta_{\mu\nu} + \kappa\psi_{\mu\nu}(X) \; , \tag{117}$$

and observe that the action (112) is related to the Polyakov action in flat space by

$$S_P[G_{\mu\nu}] = S_P[\eta_{\mu\nu}] + \kappa V[\psi_{\mu\nu}] \; , \tag{118}$$

where

$$V[\psi_{\mu\nu}] = \frac{1}{4\pi\alpha'} \int d^2\sigma\sqrt{h} h^{\alpha\beta}\psi_{\mu\nu}(X)\partial_\alpha X^\mu \partial_\beta X^\nu \; . \tag{119}$$

Taking the Fourier transform of $\psi_{\mu\nu}(X)$ we obtain

$$V[\psi_{\mu\nu}] = \frac{1}{4\pi\alpha'} \int d^D k \int d^2\sigma\sqrt{h} V_\psi(k, \tilde{\psi}(k)) \; , \tag{120}$$

where

$$V_\psi(k, \tilde{\psi}(k)) = \tilde{\psi}_{\mu\nu}(k)\partial_\alpha X^\mu \partial^\alpha X^\nu e^{\mathrm{i}k_\rho X^\rho} \tag{121}$$

is the graviton vertex operator with polarisation tensor $\tilde{\psi}_{\mu\nu}(k)$.

Thus the curved space action $S_P[G_{\mu\nu}]$ is obtained by deforming the flat space action $S_P[\eta_{\mu\nu}]$ by the graviton vertex operator. Since both actions must be conformal, $V[\psi]$ must be a so-called exactly marginal operator of the world-sheet field theory. These are the operators which generate deformations of the action while preserving conformal invariance. A necessary condition is that $V[\psi]$ must be a marginal operator, which means it has weights $(0, 0)$ with respect to the original action. Such operators have the correct weight for being added to the action and generate infinitesimal deformations which preserve conformal invariance. Note that it is not guaranteed that a marginal operator is still marginal in the infinitesimally deformed theory. Only those marginal operators which stay marginal under deformation generate finite deformations of a conformal field theory and are called exactly marginal (or truly marginal).

If the integrated vertex operator $V[\psi]$ has weights $(0, 0)$, then the vertex operator (121) must have weights $(1, 1)$. This is the condition for a vertex operator to create a physical state. The resulting conditions on momenta and polarization are

$$k^2 = 0 \; , \quad k_\mu \tilde{\psi}^{(\mu\nu)} = 0 \; , \tag{122}$$

which we now recognize as the Fourier transforms of the linearized Einstein equation. This the free part of the equations of motion for the graviton and characterizes its mass and spin.

Marginal operators are not necessarily exactly marginal. The flat space action defines a free field theory on the world-sheet, which is conformally invariant at the quantum level. Thus $V[\psi]$ is exactly marginal if and only if the curved space action $S_P[G_{\mu\nu}]$ is conformally invariant. By the beta–function analysis, this is equivalent to the full vacuum Einstein equation for the metric $G_{\mu\nu}$, plus corrections in α'. This is the full, non-linear equation of motion for the graviton string state.

In order to understand the relation between the graviton string state and the background metric even better we use (118) to relate amplitudes computed using the curved space action $S_P[G_{\mu\nu}]$ and the flat space action $S_P[\eta_{\mu\nu}]$:

$$\langle V_1 \cdots V_M \rangle_G = \langle V_1 \cdots V_M e^{V[\psi]} \rangle_\eta \tag{123}$$

The operator $e^{V[\psi]}$ generates a coherent state of gravitons in flat space. This can be seen as follows: in quantum mechanics (think of the harmonic oscillator) coherent states are defined as states with minimal Heisenberg uncertainty. They are eigenstates of annihilation operators and can be constructed by exponentiating creation operators. The resulting states are not eigenstates of the number operator but are superpositions of states with all possible occupation numbers. In (123) the role of the creation operator is played by the graviton vertex operator.

In quantum field theory, coherent states are the states corresponding to classical fields. For example, in quantum electrodynamics a classical electrodynamic field can be represented as a coherent state of photons. Similarly, in gravity a curved metric (modulo global properties) can be described as a coherent state of gravitons in the Minkowski vacuum. This is realized in the above formula, where the amplitudes in the curved background can be computed equivalently by inserting the vertex operator for a coherent state of gravitons into the flat space amplitude. This is a manifestation of background independence: though we need to pick a particular background to define our theory, other backgrounds are different states in the same theory. Since consistent backgrounds must satisfy the equations of motion, one also calls them solutions of string theory. In this terminology different background geometries are different solutions of the single underlying string theory.

3.5 Effective Actions

In the last section we have seen that the equation of motion of the metric/graviton can be obtained from an effective action. Such effective actions are very convenient, because they allow us to describe string states in terms of D-dimensional field theory. Effective actions are obtained in an expansion in α' and therefore their use is limited to scales below the string scale. But given that the string scale probably is very large, they are extremely useful to extract particle physics or gravitational physics from string theory. Therefore they play a major role in string theory. We have also seen that there are two methods for deriving effective actions: the matching of string theory amplitudes with field theory amplitudes and solving the conditions for Weyl invariance $\overline{\beta} = 0$ in a non-trivial background.

So far we found that the Einstein-Hilbert action is the leading part of the effective action for the graviton. We have seen that the closed string has two further massless modes, the dilaton Φ and the antisymmetric tensor field $B_{\mu\nu}$. We can now switch on the corresponding non-trivial background fields. The total world-sheet action is:

$$S = S_P[G] + S[B] + S[\Phi] \ . \tag{124}$$

Here $S_P[G]$ is the action (112),

$$S[B] = \frac{1}{4\pi\alpha'} \int d^2\sigma \varepsilon^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X^\nu B_{\mu\nu}(X) \tag{125}$$

and

$$S[\Phi] = \frac{1}{4\pi} \int d^2\sigma \sqrt{h} R^{(2)}(h) \Phi(X) \ . \tag{126}$$

Here $\varepsilon^{\alpha\beta}$ is the totally antisymmetric world-sheet tensor density and $R^{(2)}(h)$ is the Ricci scalar of the world-sheet metric. Note that the dilaton action is higher order in α'.

The beta-function for the dilaton starts with a term proportional to $(D-26)$ and has α'-correction proportional to derivatives of Φ. The leading term of the beta-function corresponds to a cosmological constant in the effective action. When considering string theory around backgrounds with constant dilaton, the only solution to the dilaton beta-function equation is to work in the critical dimension $D = 26$. We will only consider such backgrounds here, and therefore the cosmological term in the effective action vanishes. But let us note that there are known exact solutions to the beta-function equations with non-constant dilaton. These describe exact string backgrounds with $D \neq 26$.

Let us now return to the dilaton term of the world-sheet action. When evaluated for constant dilaton, (126) is proportional to the Euler number of the world-sheet. For a Euclidean closed string world-sheet of genus g we have:

$$\chi = \frac{1}{4\pi} \int_{\Sigma_g} d^2 z \sqrt{h} R^{(2)}(h) = 2 - 2g \ . \tag{127}$$

Therefore shifting the dilaton by a constant a,

$$\Phi(X) \rightarrow \Phi(X) + a \tag{128}$$

has the effect of shifting the total action (124) by a constant proportional to the Euler number:

$$S \rightarrow S + a\chi(g) \ . \tag{129}$$

For the corresponding partition function this is equivalent to rescaling the coupling by e^a:

$$Z = \sum_{g=0}^{\infty} \kappa^{-\chi(g)} \int DXDhe^{-S} \longrightarrow \sum_{g=0}^{\infty} (\kappa e^a)^{-\chi(g)} \int DXDhe^{-S} \ . \tag{130}$$

This shows that the coupling constant κ and vacuum expectation value $\langle \Phi \rangle$ of the dilaton are not independent. To clarify the physical meaning of both quantities, we now investigate the effective action of the massless modes. The conditions for Weyl invariance of (124) are the Euler-Lagrange equation of the following effective action:

$$S_{\text{tree}}^{\text{StrFr}} = \frac{1}{2\kappa^2} \int d^D x \sqrt{G} e^{-2\Phi} \left(R(G) - \frac{1}{12} H_{\mu\nu\rho} H^{\mu\nu\rho} + 4\partial_\mu \Phi \partial^\mu \Phi + \mathcal{O}(\alpha') \right) . \tag{131}$$

This way of parametrizing the effective action is called the string-frame. The string-frame metric $G_{\mu\nu}$ is the metric appearing in the world-sheet action (112). The field strength of the antisymmetric tensor field is

$$H_{\mu\nu\rho} = 3!\, \partial_{[\mu} B_{\nu\rho]} , \tag{132}$$

where $[\mu\nu\rho]$ denotes antisymmetrization.

Concerning the dilaton we note that its vacuum expectation value is not fixed by the equations of motion. Like in the partition function (130), shifting the dilaton by a constant is equivalent to rescaling the coupling. In order to determine the relation of the string coupling constant κ to the physical gravitational coupling κ_{phys} one has to perform a field redefinition that transforms the gravitational term in (131) into the standard Einstein-Hilbert action. The coefficient in front of this term is the physical gravitational coupling.

The transformation which achieves this is the following Weyl rescaling of the metric:

$$g_{\mu\nu} := G_{\mu\nu} e^{-\frac{4}{D-2}(\Phi - \langle \Phi \rangle)} . \tag{133}$$

Expressing everything in terms of the Einstein frame metric $g_{\mu\nu}$ one obtains:

$$\begin{aligned} S_{\text{tree}}^{\text{EinstFr}} = \frac{1}{2\kappa_{\text{phys}}^2} \int \sqrt{g} \Bigg(R(g) &- \frac{1}{12} e^{-8\frac{\Phi - \langle \Phi \rangle}{D-2}} H_{\mu\nu\rho} H^{\mu\nu\rho} \\ &- \frac{4}{D-2} \partial_\mu \Phi \partial^\mu \Phi + \mathcal{O}(\alpha') \Bigg) . \end{aligned} \tag{134}$$

The physical gravitational coupling is

$$\kappa_{\text{phys}} = \kappa e^{\langle \Phi \rangle} . \tag{135}$$

Since the coupling κ can be rescaled by shifting Φ, it can be set to an arbitrary value. This is used to fix κ:

$$\kappa \stackrel{!}{=} (\alpha')^{\frac{D-2}{4}} . \tag{136}$$

(Note that the D-dimensional gravitational couplings $\kappa, \kappa_{\text{phys}}$ have dimension $(\text{length})^{D-2/2}$.) Since κ_{phys} and α' are related by the vacuum expectation value of the dilaton we see that there is only one fundamental dimensionful parameter in string theory, which we can take to be either the gravitational coupling κ_{phys} or the string scale set by α'. They are related by the vacuum expectation value of the dilaton, which classically is a free parameter labeling different ground states in one theory. Defining the dimensionless string coupling constant by

$$g_S = e^{\langle \Phi \rangle} , \tag{137}$$

we have the relation

$$\kappa_{\text{phys}} = (\alpha')^{\frac{D-2}{4}} g_S . \tag{138}$$

The effective actions (131,134) have been constructed to leading order in α' and at tree level in the string coupling g_S. Loop corrections in g_S can be obtained, either by considering loop amplitudes or from the contribution of higher genus world-sheets to the Weyl anomaly (Fischler-Susskind mechanism). One might expect that loop corrections generate a potential for the dilaton and lift the vacuum degeneracy. But for the bosonic string one does not know the stable ground state, because of the tachyon. In supersymmetric string theories tachyons are absent, but no dilaton potential is created at any loop level. Thus the value of the string coupling remains a free parameter. This is (part of) the problem of vacuum degeneracy of superstring theories. Since the flatness of the dilaton potential is a consequence of supersymmetry, the solution of the vacuum degeneracy problem is related to understanding supersymmetry breaking.

For practical applications, both the string frame effective action and the Einstein frame effective action (and their higher-loop generalizations) are needed. The string frame action is adapted to string perturbation theory and has a universal dependence on the dilaton and on the string coupling:

$$S^{\text{StrFr}}_{\text{g}-\text{loop}} \sim g_S^{-2+2g} . \tag{139}$$

The Einstein frame action is needed when analyzing gravitational physics, in particular for solutions of the effective action that describe black holes and other space-time geometries. Note that concepts such as the ADM mass of an asymptotically flat space-time are tied to the gravitational action written in the Einstein frame. The relation between the Einstein frame metric and the string frame metric is non-trivial, because it involves the dilaton, which in general is a space-time dependent field. Therefore various quantities, most importantly the metric itself, can take a very different form in the two frames. For example one metric might be singular wheras the other is not. In order to decide whether a field configuration is singular or not, one has of course to look at all the fields, not just at the metric. If the metric is singular in one frame but not in the other, then the dilaton must be singular.

3.6 Interacting Open and Non-oriented Strings

We now indicate how the above results extend to open and non-oriented strings.

Open Strings. The world-sheets describing the interactions of open strings have two kinds of boundaries: those corresponding to the initial and final strings and those corresponding to the motions of string endpoints. Boundaries corresponding to external strings can be mapped to punctures and are then replaced by vertex operators. The boundaries corresponding to the motions of string endpoints remain. They are the new feature compared to closed strings. Perturbation theory for open strings can then be developed along the same lines as for closed strings. Instead of closed oriented surfaces it involves oriented surfaces with boundaries, and the vertex operators for open string states are inserted at on the boundaries.

Again there is one fundamental interaction, which couples three open strings, and we assign to it a coupling constant κ_o. The most simple world-sheet, analogous to the sphere for closed strings, is the disc. It is leading in the expansion in κ_o and describes scattering at tree level. The computation of tree level scattering amplitudes confirms the interpretation of the massless state as a gauge boson. The resulting effective action, to leading order in α', is the Maxwell or, with Chan-Paton factors, the Yang-Mills action. It receives higher order corrections in α' and one can show that the resulting actions are of Born-Infeld type.

Higher order diagrams in open string perturbation theory correspond to surfaces with more than one boundary component. They are obtained from the disc by removing discs from the interior. Each removal of a disc corresponds to an open string loop. The one loop diagram is the annulus.

One can also introduce a coupling of two open strings to one closed string with coupling κ_{oc} and consider theories of open and closed strings. Unitarity then implies that the three couplings $\kappa_o, \kappa_{oc}, \kappa$ are not independent. To see this consider first a disc diagram with two open string vertex operators at the boundaries and two closed string vertex operators in the interior. This amplitude can be factorized with an intermediate closed string. Looking at string interactions we see that one has one interaction between three closed strings and one between one closed and two open strings. Therefore the amplitude is proportional to $\kappa\kappa_{oc}$. The amplitude can also be factorized with an intermediate open string. This time one sees two interactions involving two open and one closed string. Therefore the amplitude is proportional to κ_{oc}^2. Comparing both forms of the amplitude we deduce

$$\kappa \simeq \kappa_{oc} \tag{140}$$

(the numerical factor has to be determined by explicit computation).

Next consider the open string one loop diagram, the annulus. Putting two vertex operators on each boundary one can again factorize it with either a closed or an open intermediate state. This way one finds

$$\kappa \simeq \kappa_o^2 \, . \tag{141}$$

Note that the above amplitude does not involve external closed string states. This indicates an important property of open string theories: the coupling to closed strings is not optional, but mandatory. When computing open string loop diagrams, one finds that they have poles which correspond to closed string states. Therefore consistency of open string theories at the quantum level requires the inclusion of closed strings. This means in particular that every consistent quantum string theory has to include gravity. The relation between open and closed strings becomes obvious when one realizes that the annulus is topologically equivalent to the cylinder. While the annulus intuitively is the open string one loop diagram, the cylinder is the closed string propagator. This is reflected by the properties of the corresponding string amplitudes, which can be written either as a sum over poles corresponding to open strings (open string channel) or as a sum over poles corresponding to closed strings (closed string channel).

The UV finiteness of closed string theories is due to modular invariance. Open string world-sheets do not have a modular group. The role of modular invariance is played by another property, called tadpole cancellation. The underlying observation is that the cancellation of divergencies between different diagrams is equivalent to the vanishing of the dilaton tadpole. It turns out that tadpole cancellation cannot be realized in a theory of oriented open and closed strings. In theories of non-oriented open and closed strings tadpole cancellation fixes the gauge group to be $SO(2^{D/2})$. For bosonic strings the critical dimension is $D = 26$ and the gauge group must be $SO(8192)$. Since the primary problem of bosonic strings is the tachyon, it is not clear whether tadpole cancellation plays a fundamental role there. But for type I superstrings this is the condition which makes the theory finite.

Since we only discussed orientable world-sheets so far, we next collect some properties of the world-sheets of non-oriented strings.

Non-oriented Strings. Theories of non-oriented strings are obtained by keeping only states which are invariant under world-sheet parity. Since such theories are insensitive to the orientation of the world–sheet one now has to include non-orientable world-sheets. Theories of closed non-oriented strings involve orientable and non-orientable surfaces without boundaries, whereas theories of open and closed non-oriented strings involve in addition orientable and non-orientable world-sheets with boundaries. Let us summarize which types of world-sheets occur in string theory, depending on boundary conditions and orientability of the world-sheet:

Strings				Surfaces			
				boundaries		orientable	
open	closed	oriented	non-oriented	without	with	yes	no
x	–	x	–	x	–	x	–
x	x	x	–	x	x	x	–
x	–	–	x	x	–	x	x
x	x	–	x	x	x	x	x

(142)

The simplest example of a non-orientable surface without boundary is the real projective plane $\mathbf{RP}^2$, which is obtained from $\mathbf{R}^2$ by adding a circle at infinity, such that every line through the origin in $\mathbf{R}^2$ intersects the circle in one point. Equivalently, $\mathbf{RP}^2$ is obtained from the disc by identifying antipodal points on its boundary. Thus, $\mathbf{RP}^2$ is a closed, but non-orientable surface, and it is a world-sheet occuring in theories of closed non-oriented strings. It is useful to note that $\mathbf{RP}^2$ can be obtained from the sphere, which is the tree-level world-sheet already familiar from oriented closed strings, by the following procedure: start with the sphere, remove a disc, (realize that the result is a disc itself,) then identify antipodal points on the resulting boundary. This operation is called 'adding a crosscap'. By iterating this process we get an infinite series of new non-orientable surfaces. For example, by adding a second crosscap we get the Klein bottle. As we discussed above, there is a similar operation that generates all orientable closed surfaces from the sphere: adding a handle. By adding both handles and crosscaps we can generate all closed surfaces, orientable and non-orientable. In fact, it is sufficient to either add handles (generating all orientable surfaces) or to add crosscaps (generating all non-orientable surfaces). The reason is that adding a crosscap and a handle is equivalent to adding three crosscaps.

When considering theories of non-oriented open strings one has to add world-sheets with boundaries. These are obtained from the world-sheets of closed strings by removing discs. For example, removing one disc from $\mathbf{RP}^2$ gives the Möbius strip. As we discussed in the last section, the couplings between open strings, κ_o, and between open and closed strings, κ_{og}, are related to the closed string coupling κ by unitarity. The order of a given world-sheet in string perturbation theory is $\kappa^{-\chi(g,b,c)}$, where the Euler number is now determined by the number g of handles, the number b of boundary components and the number c of crosscaps:

$$\chi(g,b,c) = 2 - 2g - b - c\ . \tag{143}$$

Let us write down explicitly the first few world-sheets:

g	b	c	χ	Surface	Coupling
0	0	0	2	Sphere	κ^{-2}
0	1	0	1	Disc	κ^{-1}
0	0	1	1	Real projective plane	κ^{-1}
1	0	0	0	Torus	κ^{0}
0	0	2	0	Klein bottle	κ^{0}
0	2	0	0	Annulus = cylinder	κ^{0}
0	1	1	0	Möbius strip	κ^{0}

(144)

3.7 Further Reading

Vertex operators and the Polyakov path integral are discussed in all the standard textbooks [1–4]. A very nice introduction to the use of conformal field theory in string theory is provided by [6]. A more detailed introduction to the Polyakov path integral can be found in [8]. For an extensive review of this subject, see [7]. A pedagogical treatment of the mathematical ingredients needed to treat higher genus surfaces can be found in [10], whereas [9] discusses the Polyakov path integral from the mathematicians point of view.

4 Supersymmetric Strings

The bosonic string does not have fermionic states and therefore it cannot be used as a unified theory of particle physics and gravity. One way to introduce fermionic states is an extension known as the Ramond-Neveu-Schwarz string (RNS string). In this model one introduces new dynamical fields $\psi^\mu = (\psi^\mu_A)$ on the world-sheets, which are vectors with respect to space-time but spinors with respect to the world-sheet. We will suppress the world-sheet spinor indices $A = 1, 2$ most of the time. Surprisingly, the presence of such fields, when combined with a certain choice of boundary conditions, yields states which are spinors with respect to space-time, as we will see below.

The RNS model contains space-time bosons and fermions, but still has a tachyonic ground state. One then observes that there are projections of the spectrum which simultanously remove the tachyon and make the theories space-time supersymmetric. A closer inspection shows that these projections are not optional, but required by consistency at the quantum level. This way one obtains three consistent supersymmetric strings theories, called type I, type IIA and type IIB. Finally there are also two so-called heterotic string theories, which are the result of a hybrid construction, combining type II and bosonic strings. This makes a total of five supersymmetric string theories.

4.1 The RNS Model

We now discuss the classical and quantum properties of the RNS string, proceeding along the same lines as we did for the bosonic string.

The RNS Action. The action of the RNS model is obtained from the Polyakov action by extending it to an action with supersymmetry on the world-sheet. Note that world-sheet supersymmetry is different from, and does not imply, supersymmetry in space-time. The action of the RNS model is constructed by extending the Polyakov action (4) to an action with local world-sheet supersymmetry. This action also has local Weyl symmetry, and further local fermionic symmetries which make it locally superconformal. We will not need its explicit form here. The analogue of the conformal gauge is called superconformal gauge. In this gauge the action reduces to

$$S_{\mathrm{RNS}} = \frac{1}{4\pi\alpha'} \int_{\Sigma} d^2\sigma \left(\partial_\alpha X^\mu \partial^\alpha X_\mu + \mathrm{i}\, \overline{\psi}^\mu \rho^\alpha \partial_\alpha \psi_\mu \right) . \tag{145}$$

The fields $\psi^\mu = (\psi^\mu_A)$ are Majorana spinors with respect to the world-sheet and vectors with respect to space-time, while $\rho^\alpha = (\rho^\alpha_{AB})$ are the two-dimensional spin matrices. We will usually suppress the world-sheet spinor index $A, B = 1, 2$. The action (145) is invariant under global world-sheet supersymmetry transformations:

$$\delta X^\mu = \overline{\varepsilon} \psi^\mu \;, \quad \delta \psi^\mu = -\mathrm{i}\, \rho^\alpha \varepsilon \partial_\alpha X^\mu \;. \tag{146}$$

The equations of motion are:

$$\partial^2 X^\mu = 0 \;, \quad \rho^\alpha \partial_\alpha \psi^\mu = 0 \;. \tag{147}$$

To these one has to add the constraints, which arise from the locally superconformal action. In this action the supersymmetric partner of the world-sheet metric is a vector-spinor, the gravitino. This field is non-dynamical in two dimensions and is set to zero in the superconformal gauge. The equation of motion for the metric implies that the energy-momentum tensor vanishes on shell:

$$T_{\alpha\beta} = \partial_\alpha X^\mu \partial_\beta X_\mu + \frac{i}{2} \overline{\psi}^\mu \rho_{(\alpha} \partial_{\beta)} \psi_\mu - \text{ Trace } = 0 \;. \tag{148}$$

The equation of motion for the gravitino implies that the world-sheet supercurrent J_α vanishes on shell:

$$J_\alpha = \frac{1}{2} \rho^\beta \rho_\alpha \psi^\mu \partial_\beta X_\mu = 0 \;. \tag{149}$$

In order to solve the equation of motion for ψ^μ it is convenient to choose the following spin matrices:

$$\rho^0 = \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \;, \quad \rho^1 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \;. \tag{150}$$

Using the chirality matrix $\overline{\rho} = \rho^0 \rho^1$ we see that the components $\psi^\mu_\pm$ of ψ^μ, defined by

$$\psi^\mu = \begin{pmatrix} \psi_- \\ \psi_+ \end{pmatrix} \tag{151}$$

with respect to the basis (150) are Majorana-Weyl spinors. The equations of motion decouple,

$$\partial_- \psi^\mu_+ = 0 \;, \quad \partial_+ \psi^\mu_- = 0 \;, \tag{152}$$

and have the general solution

$$\psi_+^\mu = \psi_+^\mu(\sigma^+) \ , \quad \psi_-^\mu = \psi_-^\mu(\sigma^-) \ . \tag{153}$$

Next we have to specify the boundary conditions. Requiring the vanishing of the boundary terms when varying the action implies:

$$(\psi_-^\mu \delta\psi_{\mu-} - \psi_+^\mu \delta\psi_{\mu+})\big|_{\sigma^1=0} = (\psi_-^\mu \delta\psi_-^\mu - \psi_+^\mu \delta\psi_{\mu+})\big|_{\sigma^1=\pi} \ . \tag{154}$$

For open strings we take

$$\psi_+^\mu(\sigma^0, \sigma^1 = 0) = \psi_-^\mu(\sigma^0, \sigma^1 = 0) \tag{155}$$

$$\psi_+^\mu(\sigma^0, \sigma^1 = \pi) = \pm\psi_-^\mu(\sigma^0, \sigma^1 = \pi) \ . \tag{156}$$

This couples ψ_+^μ and ψ_-^μ at the boundaries. Depending on the choice of sign in (156) one gets Ramond boundary conditions ('+' sign) or Neveu-Schwarz boundary conditions ('−' sign). One can use the same doubling trick that we used to obtain the Fourier expansion for bosonic open strings. Setting

$$\psi^\mu(\sigma^0, \sigma^1) := \begin{cases} \psi_-^\mu(\sigma^0, -\sigma^1) \text{ if} - \pi \leq \sigma^1 \leq 0 \ , \\ \psi_+^\mu(\sigma^0, \sigma^1) \quad \text{if } \ 0 \leq \sigma^1 \leq \pi \ , \end{cases} \tag{157}$$

we find that ψ is periodic for R(amond)-boundary conditions and antiperiodic for N(eveu-)S(chwarz)-boundary conditions on the doubled world-sheet. Consistency at the loop level requires that both types of boundary conditions have to be included. The Hilbert space has both an NS-sector and an R-sector.

For closed strings we can make ψ_+ and ψ_- either periodic (R-boundary conditions) or antiperiodic (NS-boundary conditions):

$$\psi_+^\mu(\sigma^0, \sigma^1 = \pi) = \pm\psi_+^\mu(\sigma^0, \sigma^1 = 0) \ , \tag{158}$$

$$\psi_-^\mu(\sigma^0, \sigma^1 = \pi) = \pm\psi_-^\mu(\sigma^0, \sigma^1 = 0) \ . \tag{159}$$

Since ψ_+^μ and ψ_-^μ are independent, one has four different choices of fermionic boundary conditions: R-R, NS-R, R-NS, NS-NS. Again considerations at the loop level require that all four sectors have to be included.

We can now write down solutions of (152) subject to the boundary conditions that we admit. For open strings we use the doubling trick and Fourier expand (157). For R-boundary conditions one obtains,

$$\psi_\mp^\mu = \frac{1}{\sqrt{2}} \sum_{n \in \mathbf{Z}} d_n^\mu e^{-in\sigma^\mp} \ , \tag{160}$$

while for NS-boundary conditions the result is:

$$\psi_\mp^\mu = \frac{1}{\sqrt{2}} \sum_{r \in \mathbf{Z}+\frac{1}{2}} b_r^\mu e^{-ir\sigma^\mp} \ . \tag{161}$$

For closed strings R-boundary conditions in the right-moving sector we get:

$$\psi^\mu_- = \sum_{n \in \mathbf{Z}} d^\mu_n e^{-2in\sigma^-} \ , \tag{162}$$

while with NS-boundary conditions this becomes

$$\psi^\mu_- = \sum_{r \in \mathbf{Z}+\frac{1}{2}} b^\mu_r e^{-2ir\sigma^-} \ . \tag{163}$$

The Fourier coefficients of the left-moving fields are denoted $\tilde{d}^\mu_n$ and $\tilde{b}^\mu_r$, respectively.

Likewise, one obtains Fourier coefficients of the energy momentum tensor $T_{\alpha\beta}$ and of the supercurrent J_α. For open strings the Fourier coefficients of J_+, J_- (in the doubled intervall) are denoted F_m in the R-sector and G_r in the NS-sector. For closed strings the Fourier modes of J_+ are denoted F_m, G_r, while those of J_- are $\tilde{F}_m$ and $\tilde{G}_r$. The Fourier components of T_{++} and T_{--} are denoted as before.

Covariant Quantization of the RNS Model. The covariant quantization of the RNS model proceeds along the lines of the bosonic string. We will consider open strings for definiteness. The canonical commutation relations of the α^μ_m are as before. The fermionic modes satisfy the canonical anticommutation relations

$$\{b^\mu_r, b^\nu_s\} = \eta^{\mu\nu}\delta_{r+s,0} \tag{164}$$

in the NS-sector and

$$\{d^\mu_m, d^\nu_n\} = \eta^{\mu\nu}\delta_{m+n,0} \tag{165}$$

in the R-sector. (For closed strings there are analogous relations for the second set of of modes.)

The Virasoro generators get contributions from both the bosonic and the fermionic oscillators, $L_m = L^{(\alpha)}_m + L^{(NS)/(R)}_m$. The bosonic part $L^{(\alpha)}_m$ is given by (29), while the contributions from the fermionic oscillators in the respective sectors are:

$$L^{(NS)}_m = \frac{1}{2}\sum_{r=-\infty}^{\infty} (r + \frac{1}{2}m)\, b_r \cdot b_{m+r} \ , \tag{166}$$

$$L^{(R)}_m = \frac{1}{2}\sum_{n=-\infty}^{\infty} (n + \frac{1}{2}m)\, d_n \cdot d_{m+n} \ . \tag{167}$$

The explicit formulae for the modes of the supercurrent are:

$$G_r = \sum_{n=-\infty}^{\infty} \alpha_{-n} \cdot b_{r+n} \ , \tag{168}$$

$$F_m = \sum_{n=-\infty}^{\infty} \alpha_{-n} \cdot d_{m+n} \ . \tag{169}$$

The modes of $T_{\alpha\beta}$ and J_α generate a supersymmetric extension of the Virasoro algebra. In the NS sector this algebra takes the form

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{D}{8}(m^3 - m)\delta_{m+n,0} \,, \tag{170}$$

$$[L_m, G_r] = (\frac{1}{2}m - r)G_{m+r} \,, \tag{171}$$

$$\{G_r, G_s\} = 2L_{r+s} + \frac{D}{2}(r^2 - \frac{1}{4})\delta_{r+s,0} \,, \tag{172}$$

while in the R-sector one finds

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{D}{8}m^3\delta_{m+n,0} \,, \tag{173}$$

$$[L_m, F_n] = (\frac{1}{2}m - n)F_{m+n} \,, \tag{174}$$

$$\{F_m, F_n\} = 2L_{m+n} + \frac{D}{2}m^2\delta_{m+n,0} \,. \tag{175}$$

The subspace of physical states $\mathcal{F}_{\text{phys}} \subset \mathcal{F}$ is found by imposing the corresponding super Virasoro constraints. In the NS-sector the constraints are:

$$\begin{aligned} L_n|\Phi\rangle &= 0\,, \quad n > 0\,, \\ (L_0 - a)|\Phi\rangle &= 0\,, \\ G_r|\Phi\rangle &= 0\,, \quad r > 0\,. \end{aligned} \tag{176}$$

Absence of negative norm states is achieved for

$$D = 10 \text{ and } a = \frac{1}{2}\,. \tag{177}$$

(Like for bosonic strings there is the option to have a non-critical string theory with $D < 10$, which we will not discuss here.) Thus the critical dimension has been reduced to 10.

In the R-sector the constraints are:

$$\begin{aligned} L_n|\Phi\rangle &= 0\,, \quad n > 0\,, \\ (L_0 - a)|\Phi\rangle &= 0\,, \\ F_n|\Phi\rangle &= 0\,, \quad n \geq 0\,. \end{aligned} \tag{178}$$

Note that there is no normal ordering ambiguity in F_0. Since $F_0^2 = L_0$ we conclude $a = 0$. The critical dimension is 10, as in the NS-sector:

$$D = 10 \text{ and } a = 0\,. \tag{179}$$

Let us construct explicitly the lowest states of the open string in both sectors. In the NS-sector the basic momentum eigenstates satisfy

$$\alpha_m^\mu|k\rangle = 0\,, \quad m > 0\,, \tag{180}$$

$$b_r^\mu|k\rangle = 0\,, \quad r > 0 \tag{181}$$

and the constraint $(L_0 - \frac{1}{2})|\Phi\rangle = 0$ provides the mass formula:

$$\alpha' M^2 = N - \frac{1}{2} \,, \tag{182}$$

where we reinstated α'. The number operator gets an additional term $N^{(b)}$ compared to (39), which counts fermionic oscillations:

$$N^{(d)} = \sum_{r=1/2}^{\infty} r\, b_{-r} \cdot b_r \,, \tag{183}$$

$$[N, b^{\mu}_{-r}] = r\, b^{\mu}_{-r} \,. \tag{184}$$

Now we can list the states:

Occupation	Mass	State
$N = 0$	$\alpha' M^2 = -\frac{1}{2}$	$\|k\rangle$
$N = \frac{1}{2}$	$\alpha' M^2 = 0$	$b^{\mu}_{-1/2}\|k\rangle$
$N = 1$	$\alpha' M^2 = \frac{1}{2}$	$b^{\mu}_{-1/2} b^{\nu}_{-1/2}\|k\rangle$ $\alpha^{\mu}_{-1}\|k\rangle$
$N = \frac{3}{2}$	$\alpha' M^2 = 1$	$b^{\mu}_{-1/2} b^{\nu}_{-1/2} b^{\rho}_{-1/2}\|k\rangle$ $\alpha^{\mu}_{-1} b^{\nu}_{-1/2}\|k\rangle$ $b^{\mu}_{-3/2}\|k\rangle$

(185)

Thus the NS-sector of the open string consists of space-time bosons and has a tachyonic ground state. The massless state is a gauge boson.

The basic momentum eigenstates in the R-sector are defined by:

$$\alpha^{\mu}_{m}|k\rangle = 0 \,, \quad m > 0 \,, \tag{186}$$

$$d^{\mu}_{m}|k\rangle = 0 \,, \quad m > 0 \,. \tag{187}$$

The constraint $L_0|\Phi\rangle = 0$ yields the mass formula

$$\alpha' M^2 = N \,. \tag{188}$$

The number operator gets an additional fermonic contribution

$$N^{(d)} = \sum_{m=1}^{\infty} m\, d_{-m} \cdot d_m \,. \tag{189}$$

The zero modes d^{μ}_0 of the fermionic fields play a distinguished role. Their algebra is, up to normalization, the Clifford algebra $Cl(1,9)$:

$$\{d^{\mu}_0, d^{\nu}_0\} = \eta^{\mu\nu} \,. \tag{190}$$

The unique irreducible representation of this algebra is the spinor representation of the Lorentz group $SO(1,9)$. Introducing standard Clifford generators $\gamma^\mu = \sqrt{2}d_0^\mu$, the generators of the spinor representation are $\sigma^{\mu\nu} = \frac{1}{4}[\gamma^\mu, \gamma^\nu]$. Since the d_0^μ are real, this representation is the 32-dimensional Majorana representation, denoted [32].

The zero modes d_0^μ commute with the number operator. Therefore the states in the R-sector organize themselves into spinor representations of the Lorentz group. This is how space-time spinors are described in the RNS model. To construct the states, we denote the ground state of the R-sector by

$$|a\rangle \ , \quad a = 1, \dots, 32 = 2^{D/2} \ , \tag{191}$$

where a transforms in the [32] representation. Then the first states are:

Occupation	Mass	State
$N=0$	$\alpha' M^2 = 0$	$\|a\rangle$
$N=1$	$\alpha' M^2 = 1$	$d_{-1}^\mu \|a\rangle$
		$\alpha_{-1}^\mu \|a\rangle$

(192)

The constraints $L_n|\Phi\rangle = 0$ $(n > 0)$ and the new constraints $F_n|\Phi\rangle = 0$ $(n \geq 0)$ impose restrictions on the polarization. For example, $F_0|a\rangle = 0$ is easily seen to be the Fourier transform of the massless Dirac equation and reduces the number of independent components by a factor $\frac{1}{2}$. Excited states are obtained by acting with creation operators α_{-m}^μ, d_{-m}^μ on the gound state. Since the product of a tensor representation of the Lorentz group with a spinor representation always gives spinor representations, we see that all states in the R-sector are space-time spinors.

The GSO Projection for Open Strings. The RNS model solves the problem of describing space-time fermions but still has a tachyon. Gliozzi, Scherk and Olive observed that one can make a projection of the spectrum, which removes the tachyon. Moreover, the resulting spectrum is supersymmetric in the space-time sense. This so-called GSO projection is optional at the classical level, but it becomes mandatory at the quantum level, as we will discuss below.

The GSO projector in the NS-sector is defined as follows:

$$P_{GSO}^{(NS)} = -(-1)^{\sum_{r=1/2}^{\infty} b_{-r}\cdot b_r} \ . \tag{193}$$

Imposing $P_{GSO}^{(NS)} \stackrel{!}{=} 1$, one projects out all the states which contain an even number of b^μ_{-r} creation operators. This in particular removes the tachyon. The GSO projector in the R-sector is

$$P^{(R)}_{GSO} = \overline{\gamma}(-1)^{\sum_{m=1}^{\infty} d_{-m}\cdot d_m} , \tag{194}$$

where $\overline{\gamma}$ is the ten-dimensional chirality operator. On the ground state $|a\rangle$ of the R-sector the projection $P^{(R)}_{GSO}|\Phi\rangle \stackrel{!}{=} 1$ removes one chirality of the spinor. This is consistent, because in ten space-time dimensions the irreducible spinor representations are Majorana-Weyl spinors. The [32] representation decomposes according to

$$[32] = [16]_+ + [16]_- . \tag{195}$$

With the GSO projection one only keeps one chirality (which we have taken to be the $[16]_+$, for definiteness):

$$|a\rangle = |a_+\rangle + |a_-\rangle \longrightarrow |a_+\rangle , \tag{196}$$

where $a_+ = 1, \ldots , 16$ is a Majorana-Weyl index.

At the massive level just projecting out one chirality would not be consistent, as massive particles cannot be chiral. The projection with (194) keeps states which either have '+' chirality and an even number of d^μ_{-m} creation operators or '−' chirality and an odd number of d^μ_{-m} creation operators.

By writing down the first few states one can easily verify that after the projection the NS-sector and R-sector have an equal number of states, and that the massive states in the R-sector combine into full (non-chiral) massive Lorentz representations.

Checking the equality of states at every mass level is done by computing the one-loop partition function. Moreover, one can construct explicitly the representation of the ten-dimensional super Poincaré algebra on the physical states. This is done using BRST techniques and lies beyond the scope of these lectures. Here we restrict ourselves to noting that the ground state of the open string, after GSO projection, is a ten-dimensional vector supermultiplet:

$$\{b^\mu_{-1/2}|k\rangle , \quad |a_+\rangle\} . \tag{197}$$

Spectrum and GSO Projection for Closed Strings. Let us next study the spectrum of closed RNS strings. The masses of states are determined by

$$\begin{aligned} \alpha' M^2 &= 2(N - a_x + \tilde{N} - \tilde{a}_x) , \\ N - a_x &= \tilde{N} - \tilde{a}_x , \end{aligned} \tag{198}$$

with normal ordering constants $a_R = 0 = \tilde{a}_R$ and $a_{NS} = \frac{1}{2} = \tilde{a}_{NS}$.

We start by listing the first states in the NS-NS sector:

Occupation	Mass	States
$N = \tilde{N} = 0$	$\alpha' M^2 = -2$	$\|k\rangle$
$N = \tilde{N} = \frac{1}{2}$	$\alpha' M^2 = 0$	$b^\mu_{-1/2} \tilde{b}^\nu_{-1/2} \|k\rangle$
$N = \tilde{N} = 1$	$\alpha' M^2 = 2$	$\alpha^\mu_{-1} \tilde{\alpha}^\nu_{-1} \|k\rangle$
		$\alpha^\mu_{-1} \tilde{b}^\nu_{-1/2} \tilde{b}^\rho_{-1/2} \|k\rangle$
		$b^\mu_{-1/2} b^\nu_{-1/2} \tilde{\alpha}^\rho_{-1} \|k\rangle$
		$b^\mu_{-1/2} b^\nu_{-1/2} \tilde{b}^\rho_{-1/2} \tilde{b}^\sigma_{-1/2} \|k\rangle$

(199)

All these states are bosons, and at the massless level we recognize the graviton, the dilaton and the antisymmetric tensor.

In the R-R sector, the ground state transforms in the $[32] \times [32]$ representation and is denoted $|a, \tilde{a}\rangle$. The first states are

Occupation	Mass	State
$N = \tilde{N} = 0$	$\alpha' M^2 = 0$	$\|a, \tilde{a}\rangle$
$N = \tilde{N} = 1$	$\alpha' M^2 = 2$	$\alpha^\mu_{-1} \tilde{\alpha}^\nu_{-1} \|a, \tilde{a}\rangle$
		$d^\mu_{-1} \tilde{\alpha}^\nu_{-1} \|a, \tilde{a}\rangle$
		$\alpha^\mu_{-1} \tilde{d}^\nu_{-1} \|a, \tilde{a}\rangle$
		$d^\mu_{-1} \tilde{d}^\nu_{-1} \|a, \tilde{a}\rangle$

(200)

The product of two spinor representations is a vector-like representation. Therefore the states in the R-R sector are bosons. In more detail, the $[32] \times [32]$ representation is the direct sum of all the antisymmetric tensor representations of rank zero to ten. Using the ten-dimensional Γ-matrices we can decompose a general massless state into irreducible representations:

$$|\Phi_{RR}\rangle = (F \delta_{a\tilde{a}} + F_\mu \Gamma^\mu_{a\tilde{a}} + F_{\mu\nu} \Gamma^{\mu\nu}_{a\tilde{a}} + \cdots)|a, \tilde{a}\rangle \; . \tag{201}$$

By evaluating the remaining constraints $F_0 |\Phi_{RR}\rangle = 0 = \tilde{F}_0 |\Phi_{RR}\rangle$ one obtains the conditions

$$k^{\mu_1} F_{\mu_1 \mu_2 \ldots \mu_n} = 0 \text{ and } k_{[\mu_0} F_{\mu_1 \mu_2 \ldots \mu_n]} = 0 \; , \tag{202}$$

which are the Fourier transforms of the equation of motion and Bianchi identity of an n-form field strength:

$$d \star F_{(n)} = 0 \text{ and } dF_{(n)} = 0 \; . \tag{203}$$

The physical fields are antisymmetric tensor gauge fields or rank $n-1$. Note that in contrast to the antisymmetric NS-NS field, the states in the R-R sector (and the corresponding vertex operators) describe the field strength and not the gauge potential. When analyzing interactions one finds that there are no minimal gauge couplings but only momentum couplings of these fields (i.e. couplings involving the field strength). In other words the perturbative spectrum does not contain states which are charged under these gauge fields. This is surprising, but a closer analysis shows that the theory has solitonic solutions which carry R-R charge. These so called R-R charged p-branes turn out to be an alternative description of D-branes.

Now we turn to the NS-R sector. The first states are:

Occupation	Mass	State
$N = \frac{1}{2}\,, \tilde{N} = 0$	$\alpha' M^2 = 0$	$b^{\mu}_{-1/2} \vert\tilde{a}\rangle$
$N = \frac{3}{2}\,, \tilde{N} = 1$	$\alpha' M^2 = 4$	$\alpha^{\mu}_{-1} b^{\nu}_{-1/2} \tilde{\alpha}^{\rho}_{-1} \vert\tilde{a}\rangle$
		$b^{\mu}_{-1/2} b^{\nu}_{-1/2} b^{\rho}_{-1/2} \tilde{\alpha}^{\sigma}_{-1} \vert\tilde{a}\rangle$
		$\alpha^{\mu}_{-1} b^{\nu}_{-1/2} \tilde{d}^{\rho}_{-1} \vert\tilde{a}\rangle$
		$b^{\mu}_{-1/2} b^{\nu}_{-1/2} b^{\rho}_{-1/2} \tilde{d}^{\mu}_{-1} \vert\tilde{a}\rangle$

(204)

The massless state is a product of a vector $[D]$ and a spinor $[2^{D/2}]$. It decomposes into a vector-spinor and a spinor:

$$[D] \times [2^{D/2}] = [(D-1)2^{D/2}] + 2^{D/2} \,. \tag{205}$$

Therefore this state and all other states in the NS-R sector are space-time fermions. The spectrum of the R-NS sector is obtained by exchanging left- and right-moving fermions.

We observe that the massless states contains two vector-spinors. The only known consistent interaction for such fields is supergravity. There these fields are called gravitini. They sit in the same supermultiplet as the graviton, they are the gauge fields of local supertransformations and couple to the conserved supercurrent. The spectrum of the closed RNS model is obviously not supersymmetric. This suggests that we have to make a projection in order to obtain consistent interactions. This brings us to the GSO projection for closed strings, which makes the spectrum supersymmetric and removes the tachyon. The GSO projection is applied both in the left-moving and in the right-moving sector. In the R-sectors one has to decide which chirality one keeps. There are two inequivalent projections of the total spectrum: one either takes opposite chiralities of the R-groundstates (type A) or the same chiralities (type B). The resulting theories are the type IIA and type IIB superstring. Let us look at their massless states. The NS-NS sectors of both theories are identical. The states

$$b^{\mu}_{-1/2} \tilde{b}^{\nu}_{-1/2} |k\rangle \tag{206}$$

are the graviton $G_{\mu\nu}$, the dilaton Φ, and the antisymmetric tensor $B_{\mu\nu}$. The number of on-shell states is $8 \cdot 8 = 64$. The ground states of the R-R sectors are:

$$|a_+, \tilde{a}_-\rangle \ \ \text{(type A)} \ , \tag{207}$$

$$|a_+, \tilde{a}_+\rangle \ \ \text{(type B)} \ . \tag{208}$$

In both cases we have $8 \cdot 8 = 64$ on-shell states. Again we can decompose these representations into irreducible antisymmetric tensors. For type IIA we get a two-form and a four-form field strength, corresponding to a one-form and a three-form potential:

$$\text{IIA}: \ \ A_\mu \ , \ \ A_{\mu\nu\rho} \ . \tag{209}$$

There is also a zero-form field strength which has no local dynamics. It corresponds to the so-called massive deformation of IIA supergravity, which is almost but not quite a cosmological constant. (In the effective action the corresponding term is a dimensionful constant multiplied by the dilaton. This is as close as one can get to a cosmological constant in ten-dimensional supergravity.)

In the IIB theory one has a one-form, a three-form and a selfdual five-form field strength. The corresponding potentials are:

$$\text{IIB}: \ \ A \ , \ \ A_{\mu\nu} \ , \ \ A_{\mu\nu\rho\sigma} \ . \tag{210}$$

The massless states in the NS-R sector and R-NS sector are:

$$\text{IIA}: \ b^\mu_{-1/2}|\tilde{a}_-\rangle \ \ \tilde{b}^\mu_{-1/2}|a_+\rangle \ , \tag{211}$$

$$\text{IIB}: \ b^\mu_{-1/2}|\tilde{a}_+\rangle \ \ \tilde{b}^\mu_{-1/2}|a_+\rangle \ , \tag{212}$$

The total number of fermionic states is 128 in both cases. The decomposition into irreducible representations gives two vector-spinors, the gravitini, and two spinors, called dilatini. For type IIA they have opposite chiralities, whereas for type IIB they have the same chiralities. The corresponding space-time fields are:

$$\begin{aligned} \text{IIA}: &\ \ \psi^\mu_+ \ , \ \ \psi^\mu_- \ , \ \ \psi_+ \ , \ \ \psi_- \ , \\ \text{IIB}: &\ \ \psi^\mu_{+(1)} \ , \ \ \psi^\mu_{+(2)} \ , \ \ \psi_{+(1)} \ , \ \ \psi_{+(2)} \ . \end{aligned} \tag{213}$$

All together we get the field content of the type IIA/B supergravity multiplet with 128 bosonic and 128 fermionic on-shell states. The IIA theory is non-chiral whereas the IIB theory is chiral. The massive spectra are of course non-chiral, and, moreover, they are identical.

4.2 Type I and Type II Superstrings

We will now begin to list all consistent supersymmetric string theories. A priori, we have the following choices: strings can be (i) open or closed, (ii) oriented or non-oriented, (iii) one can make the GSO projection, with two inequivalent choices (type A and B) for closed strings and (iv) one can choose gauge groups for open strings: $U(n)$ for oriented and $SO(n)$ or $Usp(2n)$ for non-oriented strings.

We have already seen that not all combinations of these choices are consistent at the quantum level. Since theories of open strings have closed string poles in loop diagrams, we can either have closed or closed and open strings. The next restriction comes from modular invariance. On the higher genus world-sheets of closed oriented strings, one has to specify boundary conditions around every handle. Since modular invariance maps one set of boundary conditions to others, these choices are not independent. It turns out that one has to include both NS- and R-boundary conditions around every handle, but one has the freedom of choosing relative signs between different orbits of action of the modular group on the set of boundary conditions. There are four possible choices. Two of them correspond to the IIA and IIB superstrings. The other two choices are non-supersymmetric theories without fermions, known as type 0A and 0B, which we will not discuss here.

Type IIA and IIB are theories of oriented closed strings. Can we construct supersymmetric string theories with oriented closed and open strings? The states of the oriented open string fall into representations of the minimal $N = 1$ supersymmetry algebra in $D = 10$. This algebra has 16 supercharges, which transform as a Majorana-Weyl spinor under the Lorentz group. In ten dimensions there are two further supersymmetry algebras, called $N = 2A$ and $N = 2B$. They have 32 supercharges which either combine into two Majorana-Weyl spinors of opposite chirality (A) or into two Majorana-Weyl spinors of the same chirality (B). The states of the oriented closed string form multiplets of the $N = 2A$ or $N = 2B$ supersymmetry algebra. In particular one has two gravitini, which must couple to two independent supercurrents. Therefore oriented open and closed strings cannot be coupled in a supersymmetric way. One can also show that any such theory has divergencies, due to the presence of dilaton tadpoles.

Next we have to consider non-oriented strings. A theory of non-oriented closed strings can be obtained by projecting the type IIB theory onto states invariant under world-sheet parity. (IIA is not invariant under world-sheet parity, because the R-groundstates have opposite chirality.) This theory has divergencies, which are related to the non-vanishing of dilaton tadpole diagrams. One can also see from the space-time point of view that this theory is inconsistent: the massless states form the $N = 1$ supergravity multiplet, which is chiral. Pure $N = 1$ supergravity has a gravitational anomaly, which can only be cancelled by adding precisely 496 vector multiplets.

Therefore we have to look at theories with non-oriented closed and open strings. Tadpole cancellation precisely occurs if the gauge group is chosen to be $SO(2^{D/2}) = SO(32)$. This is one of the gauge groups for which gravitational anomalies cancel. The other anomaly-free gauge groups are $E_8 \times E_8$, $E_8 \times U(1)^{248}$ and $U(1)^{496}$, which, however, cannot be realized through Chan-Paton factors. Thus there is one supersymmetric string theory with non-oriented closed and open strings and gauge group $SO(32)$. This is the type I superstring.

Let us construct the massless spectrum of this theory. The closed string sector is obtained by projecting the IIB theory onto states invariant under world-sheet parity. Parity acts by exchanging left- and right-moving quantities:

$$\alpha^\mu_m \leftrightarrow \tilde{\alpha}^\mu_m \;, \quad b^\mu_r \leftrightarrow \tilde{b}^\mu_r \;, \quad d^\mu_m \leftrightarrow \tilde{d}^\mu_m \;, \quad |a_+\rangle \leftrightarrow |\tilde{a}_+\rangle \;. \tag{214}$$

The action on the R-R ground state is:

$$|a_+, \tilde{a}_+\rangle \leftrightarrow -|\tilde{a}_+, a_+\rangle \;. \tag{215}$$

The '$-$' sign reflects that one exchanges two fermionic states. (To make this precise one needs to construct the so-called spin fields S^a, $S^{\tilde{a}}$ which generate the R-groundstates from the NS-groundstate. This can be done in the framework of BRST quantization, which we did not introduce here.)

We can now write down the massless states of the type IIB string which are invariant under world-sheet parity and survive the projection. In the NS-NS sector we find

$$\text{NS-NS} \;:\quad \frac{1}{2}\left(b^\mu_{-1/2}\tilde{b}^\nu_{-1/2} + b^\nu_{-1/2}\tilde{b}^\mu_{-1/2}\right)|k\rangle \;. \tag{216}$$

Therefore the $B_{\mu\nu}$ field is projected out and we are left with the graviton $G_{\mu\nu}$ and dilaton Φ. In the R-R sector the invariant massless state is:

$$\text{R-R} \;:\quad \frac{1}{2}\left(|a_+, \tilde{a}_+\rangle - |\tilde{a}_+, a_+\rangle\right) \;. \tag{217}$$

Thus only the antisymmetric part of the tensor product of the two Majorana-Weyl spinors survives the projection. This corresponds to the three-form field strength $F_{\mu\nu\rho}$, as is most easily seen by computing the dimensions of the representations. Thus the two-form R-R gauge field $A_{\mu\nu}$ survives the projection.

In the NS-R and R-NS one finds the following invariant state:

$$\text{R-NS/NS-R} \;:\quad \frac{1}{2}\left(b^\mu_{-1/2}|\tilde{a}_+\rangle + \tilde{b}^\mu_{-1/2}|a_+\rangle\right) \;. \tag{218}$$

Therefore one gravitino ψ^μ_+ and one dilatino ψ^μ are kept.

In the NS-sector of the open string we get massless vectors A^i_μ, which transform in the adjoint representations of $SO(32)$: $i = 1, \ldots, \dim(\text{adj}SO(32)) = 496$. The R-sector contains massless spinors ψ^i which combine with the vectors to form vector supermultiplets.

Combining the massless states of the closed and open string sector we get the field content of $N = 1$ supergravity coupled to Super-Yang-Mills theory with gauge group $SO(32)$.

4.3 Heterotic Strings

There is yet another construction of supersymmetric string theories. It is a hybrid construction, which combines the bosonic string with the type II superstring and is called the heterotic string. The right-moving sector is taken from the type II superstring, whereas the left-moving sector is taken from the bosonic string. To get a modular invariant theory, the sixteen extra left-moving coordinates have to be identified periodically,

$$X^I \simeq X^I + w^I_{(i)} \;, \quad I = 1, \ldots, 16 \;. \tag{219}$$

The vectors $\boldsymbol{w}_{(i)} = (w^I_{(i)})$, $i = 1, \ldots, 16$ generate a sixteen-dimensional lattice Γ_{16}. Modular invariance requires that Γ_{16} is an even self-dual lattice. Modulo rotations, there are only two such lattices, the root lattice of $E_8 \times E_8$ and the lattice generated by the roots and the weights of one of the Majorana-Weyl spinor representations of $SO(32)$. Thus, there are two different heterotic string theories.

The bosonic massless states come from the NS-sector and take the form

$$\alpha^\mu_{-1} \tilde{b}^\nu_{-1/2} |k\rangle \tag{220}$$

$$\alpha^I_{-1} \tilde{b}^\nu_{-1/2} |k\rangle \tag{221}$$

$$e^{\mathrm{i} k^{(i)}_I x^I_L} \tilde{b}^\nu_{-1/2} |k\rangle \quad . \tag{222}$$

Here α^I_{-1} are the oscillators corresponding to the sixteen extra left-moving directions. The vectors $k^{(i)} = (k^{(i)}_I)$ are discrete momentum vectors in the extra dimensions. The above states are massless if the vectors $k^{(i)}$ have norm-squared two. The two lattices Γ_{16} have 480 such vectors, corresponding to the roots of $E_8 \times E_8$ and $SO(32)$, respectively. Together with the states generated by the internal oscillators one gets bosons in the adjoint representations of theses two groups. The massless fermionic states are obtained by replacing $\tilde{b}^\nu_{-1/2}|k\rangle$ by the R-ground state $|a_+\rangle$. In total one gets the $N = 1$ supergravity multiplet plus vector multiples in the adjoint representation of $E_8 \times E_8$ or $SO(32)$.

The massless sectors of the five supersymmetric string theories correspond to four different supergravity theories. The type I and the heterotic string with gauge group $SO(32)$ have the same massless spectrum, but their massive spectra and interactions are different.

Let us summarize the essential properties of the five supersymmetric string theories:

Type	open/closed?	oriented?	chiral?	supersymmetry	gauge group
I	both	no	yes	$N = 1$	$SO(32)$
II A	closed	yes	no	$N = 2A$	–
II B	closed	yes	yes	$N = 2B$	–
Heterotic	closed	yes	yes	$N = 1$	$E_8 \times E_8$
Heterotic	closed	yes	yes	$N = 1$	$SO(32)$

(223)

4.4 Further Reading

Supersymmetric string theories are discussed in all of the standard textbooks [1–5]. To prove the necessity of the GSO projection and the consistency of the heterotic string as a perturbative quantum theory one needs properties of the multiloop path integral [7]. A paedagogical treatment of the relation between the GSO projection and boundary conditions in the path integral can be found in [6].

5 p-Branes in Type II String Theories

In this section we discuss a class of solitons of the type II string theories, which turn out to be alternative descriptions of the D-branes introduced earlier.

5.1 Effective Actions of Type II String Theories

The effective actions for the massless states of type IIA/B superstring theory are the corresponding type IIA/B supergravity actions. Since we will be interested in bosonic solutions of the field equations, we will only display the bosonic parts. The effective action for the fields in the NS-NS sector is the same for both theories. Moreover it is identical to the effective action of the bosonic string:

$$S_{\mathrm{NS-NS}} = \frac{1}{2\kappa^2} \int d^{10}x \sqrt{-G} e^{-2\Phi} \left(R + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{12} H_{\mu\nu\rho} H^{\mu\nu\rho} \right) . \tag{224}$$

The R-R sectors consist of antisymmetric tensor gauge fields. For an $(n-1)$ form gauge potential $A_{(n-1)}$ with field strength $F_{(n)} = dA_{(n-1)}$ the generalized Maxwell action is

$$S \simeq -\frac{1}{2} \int F_{(n)} \wedge \star F_{(n)} = -\frac{1}{2} \int d^D x \sqrt{-G} |F_{(n)}|^2 , \tag{225}$$

where

$$|F_{(n)}|^2 := \frac{1}{n!} F_{\mu_1 \cdots \mu_n} F^{\mu_1 \cdots \mu_n} . \tag{226}$$

In the effective R-R actions one has in addition Chern-Simons terms.

In the IIA theory the R-R fields are $A_{(1)}$ and $A_{(3)}$. It is convenient to define a modified field strength

$$\tilde{F}_{(4)} = dA_{(3)} - A_{(1)} \wedge H_{(3)} , \tag{227}$$

where $H_{(3)} = dB_{(2)}$ is the field strength of the antisymmetric NS-NS tensor field. Then the R-R action is the sum of a Maxwell and a Chern-Simons term:

$$\begin{aligned} S^{\mathrm{IIA}}_{\mathrm{R-R}} = &-\frac{1}{4\kappa^2} \int d^{10}x \sqrt{-G} \left(|F_{(2)}|^2 + |\tilde{F}_{(4)}|^2 \right) \\ &-\frac{1}{4\kappa^2} \int B_{(2)} \wedge F_{(4)} \wedge F_{(4)} . \end{aligned} \tag{228}$$

The massless R-R fields of IIB string theory are $A_{(0)}$, $A_{(2)}$ and $A_{(4)}$. Again it is useful to define modified field strengths

$$\begin{aligned} \tilde{F}_{(3)} &= F_{(3)} - A_{(0)} \wedge H_{(3)} \;, \\ \tilde{F}_{(5)} &= F_{(5)} - \frac{1}{2} A_{(2)} \wedge H_{(3)} + \frac{1}{2} B_{(2)} \wedge F_{(3)} \;. \end{aligned} \tag{229}$$

Since $\tilde{F}_{(5)}$ must be selfdual, the kinetic term (225) vanishes and does not give a field equation. The simplest way out is to impose the selfduality condition only at the level of the equation of motion. Then one can use the action

$$\begin{aligned} S^{\mathrm{IIB}}_{\mathrm{R-R}} = &-\frac{1}{4\kappa^2} \int d^{10}x \sqrt{-G} \left(|F_{(1)}|^2 + |\tilde{F}_{(3)}|^2 + \frac{1}{2} |\tilde{F}_{(5)}|^2 \right) \\ &-\frac{1}{4\kappa^2} \int A_{(4)} \wedge H_{(3)} \wedge F_{(3)} \;. \end{aligned} \tag{230}$$

The correct covariant equations of motion result when varying the action and imposing selfduality of $\tilde{F}_{(5)}$ afterwards.

5.2 R-R Charged p-Brane Solutions

The type II effective actions have static solutions which are charged under the R-R gauge fields. The solution charged under $A_{(p+1)}$ has p translational isometries. From far it looks like a p-dimensional membrane and therefore one calls it a p-brane solution or just a p-brane.

For $0 \leq p \leq 2$ the solution has the following form:

$$\begin{aligned} ds^2_{\mathrm{Str}} &= H^{-1/2}(r) \left(-dt^2 + (dx^1)^2 + \cdots + (dx^p)^2 \right) \\ &\quad + H^{1/2}(r) \left((dx^{p+1})^2 + \cdots + (dx^9)^2 \right) \;, \\ F_{(p+2)} &= dH^{-1}(r) \wedge dt \wedge dx^1 \wedge \cdots \wedge dx^p \;, \\ e^{-2\Phi} &= H^{(p-3)/2}(r) \;, \end{aligned} \tag{231}$$

where

$$r^2 = (x^{p+1})^2 + \cdots + (x^9)^2 \;, \tag{232}$$

and $H(r)$ is a harmonic function of the transverse coordinates $(x^{p+1}, \ldots, x^9)$:

$$\Delta^{\perp} H = \sum_{i=p+1}^{9} \partial_i \partial_i H = 0 \;. \tag{233}$$

We require that the solution becomes asymptotically flat at transverse infinity and normalize the metric such that it approaches the standard Minkowski metric. This fixes

$$H(r) = 1 + \frac{Q_p}{r^{7-p}} \;. \tag{234}$$

Q_p measures the flux of the R-R field strength at transverse infinity. A convenient way to parametrize it is:

$$Q_p = N_p c_p \ , \quad c_p = \frac{(2\pi)^{7-p}}{(7-p)\omega_{8-p}} (\alpha')^{\frac{7-p}{2}} g_S \ . \tag{235}$$

N_p is a constant, which a priori is real, but will turn out later to be an integer. Therefore c_p is the fundamental quantum of R-R p-brane charge. ω_n is the volume of the n-dimensional unit sphere,

$$\omega_n = \frac{2\pi^{(n+1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \ . \tag{236}$$

Besides geometrical factors, Q_p contains the appropriate power of α' to give it the correct dimension. g_S is the dimensionless string coupling. Note that in the above solution for the dilaton we have subtracted the dilaton vacuum expectation value from Φ.

The metric used in this solution is the string frame metric, as indicated by the subscript. (The effective action was also given in the string frame.) Using (133) we can find the corresponding Einstein frame metric:

$$\begin{aligned} ds^2_{\text{Einst}} = & -H^{\frac{p-7}{8}}(r)\left(-dt^2 + (dx^1)^2 + \cdots + (dx^p)^2\right) \\ & +H^{\frac{p+1}{8}}(r)\left((dx^{p+1})^2 + \cdots + (dx^9)^2\right) \ . \end{aligned} \tag{237}$$

The above solution is most easily understood as a generalization of the extreme Reissner-Nordström solution of four-dimensional Einstein-Maxwell theory. Let us review its properties.

The isometry directions $t, x^1, \ldots, x^p$ are called longitudinal or world-volume directions, the others transverse directions. Since the solution has translational invariance, it has infinite mass, as long as one does not compactify the world-volume directions. However, the tension T_p (the energy per world volume) is finite. Since the solution becomes asymptotically flat in the transverse directions, the tension can be defined by a generalization of the ADM construction of general relativity. Concretely, the tension of a p-brane can be extracted from the Einstein frame metric by looking at the leading deviation from flatness:

$$g_{00} = -1 + \frac{16\pi G_N^{(D)} T_p}{(D-2)\omega_{D-2-p} r^{D-3-p}} + \cdots = -1 + \frac{16\pi G_N^{(10)} T_p}{8\omega_{8-p} r^{7-p}} + \cdots \tag{238}$$

The Schwarzschild radius r_S of the brane is:

$$r_S^{D-3-p} = \frac{16\pi G_N^{(D)} T_p}{(D-2)\omega_{D-2-p}} \ . \tag{239}$$

Since there is only one independent dimensionful constant, which we take to be α', we can express the ten-dimensional Newton constant $G_N^{(10)}$ in terms of α' and the dimensionless string coupling g_S:

$$G_N^{(10)} = 8\pi^6 (\alpha')^4 g_S^2 \ . \tag{240}$$

Since Newton's constant is related to the physical gravitational coupling by

$$8\pi G_N^{(D)} = \kappa^2_{(D),\text{phys}} \tag{241}$$

in any dimension, this corresponds to replacing the conventional choices (136, 138) by $\kappa^2 \stackrel{!}{=} 64\pi^7(\alpha')^4$ and $\kappa^2_{\text{phys}} = 64\pi^7(\alpha')^4 g_S^2$.

Using (238) we can compute the tension of the p-brane solution (231):

$$T_p = \frac{N_p}{g_S(\alpha')^{\frac{p+1}{2}}(2\pi)^p} \, . \tag{242}$$

For $r \to 0$ the solution (231) has a null singularity, that is a curvature singularity which is lightlike and coincides with an event horizon. The p-brane (231) is the extremal limit of a more general black p-brane solution, which has a time-like singularity along a p-dimensional surface and a regular event horizon. In the extremal limit, the singularity and the even horizon coincide. This behaviour is similar to the Reissner-Nordström black hole. The behaviour of the black p-brane in the extremal limit is slightly more singular, because for the extremal Reissner-Nordström black hole singularity and horizon do not coinicide. But since the singularity of the p-brane solution is not naked, we can think about it as describing an extended charged black hole. The charge (density) carried under the gauge field $A_{(p+1)}$ can be read off from the asymptotic behaviour of the field strength,

$$F_{01\ldots p} \simeq \frac{Q_p}{r^{8-p}} \, . \tag{243}$$

Instead of Q_p we can use a redefined charge, which has the dimension of a tension:

$$\hat{Q}_p = \frac{1}{2\kappa^2} \oint_{S_{8-p}} \star F_{(p+2)} \, , \tag{244}$$

which gives

$$\hat{Q}_p = N_p \frac{\mu_p}{g_S} \, , \quad \mu_p = \frac{1}{(2\pi)^p (\alpha')^{\frac{p+1}{2}}} \, . \tag{245}$$

We now observe that tension and charge are equal:

$$T_p = \hat{Q}_p \, . \tag{246}$$

More generally, black p-brane solutions satisfy the Bogomol'nyi bound

$$T_p \geq \hat{Q}_p \, . \tag{247}$$

This inequality guarantees the existence of an event horizon, just as for charged black holes.

A feature that distinguishes our solutions from Reissner-Nordström type black holes is that one also has a non-trivial scalar, the dilaton.

The extremal solution has a multicentered generalization. When replacing $H(r)$ by

$$H(\boldsymbol{x}_\perp) = 1 + \sum_{i=1}^{N} \frac{|Q_p^{(i)}|}{|\boldsymbol{x}_\perp - \boldsymbol{x}_\perp^{(i)}|^{7-p}} , \tag{248}$$

one still has a static solution, provided that all the charges $Q_p^{(i)}$ have the same sign. Here $\boldsymbol{x}_\perp = (x^{p+1}, \ldots , x^9)$ and $\boldsymbol{x}_\perp^{(i)}$ is the position of (the horizon of) the i-th p-brane. It is remarkable that the solution is static for arbitrary positions $\boldsymbol{x}_\perp^{(i)}$, because this implies that the gravitational attraction and the 'electrostatic' repulsion cancel for arbitrary distances. (If one flips the sign of one charge, one has to flip the corresponding tension, which makes the solution unphysical.) Systems of extremal Reissner-Nordström black holes have the same properties. The corresponding multi-centered solutions are known as Majumdar-Papapetrou solutions.

The remarkable properties of these (and other related) solutions can be understood in terms of supersymmetry. The solution (231) is a supersymmetric solution, i.e., it has Killing spinors. Killing spinors are the supersymmetric analogues of Killing vectors $v(x)$, which satisfy

$$\mathcal{L}_{v(x)} \, \Psi(x)|_{\Psi_0(x)} = 0 , \tag{249}$$

where $\mathcal{L}$ is the Lie derivative. Here $\Psi(x)$ collectively denotes all the fields, and $\Psi_0(x)$ is the particular field configuration, which is invariant under the transformation generated by the vector field $v(x)$. In supergravity theories one can look for field configurations $\Psi_0(x)$ which are invariant under supersymmetry transformations. From the action one knows the supersymmetry variations of all the fields, $\delta_{\varepsilon(x)}\Psi(x)$, where the spinor (field) $\varepsilon(x)$ is the transformation parameter. Then one can plug in a given field configuration $\Psi_0(x)$ and check whether the variation vanishes for a specific choice of $\varepsilon(x)$:

$$\delta_{\varepsilon(x)} \, \Psi(x)|_{\Psi_0(x)} = 0 . \tag{250}$$

Since the supersymmetry transformations involve derivatives of $\varepsilon(x)$, this is a system of first order differential equations for $\varepsilon(x)$. Solutions of (250) are called Killing spinors.

The type II superalgebras have 32 independent real transformation parameters, which organize themselves into two Majorana-Weyl spinors $\varepsilon_i(x)$. The equation (250) fixes the space-time dependence of the $\varepsilon_i(x)$. For the p-brane one finds

$$\varepsilon_i(x) = g_{tt}^{1/4}(x)\varepsilon_i^{(0)} , \tag{251}$$

where the constant Majorana-Weyl spinors $\varepsilon_i^{(0)}$, $i = 1, 2$ are related by

$$\varepsilon_2^{(0)} = \Gamma^0 \cdots \Gamma^p \varepsilon_1^{(0)} . \tag{252}$$

Since half of the components of the $\varepsilon_i^{(0)}$ is fixed in terms of the other half, we see that we have 16 independent solutions, i.e., 16 Killing spinors. The maximal number of Killing spinors equals the number of sypersymmetry transformation parameters, which is 32 in type II theory. Solutions with the maximal number of Killing spinors are invariant under all supersymmetry transformations. They are the analogues of maximally symmetric spaces in Riemannian geometry, which by definition have as many isometries as flat space. One example of a maximally supersymmetric solution of type II theory is flat ten-dimensional Minkowski space. Here the Killing spinor equation is solved by all constant spinors. The p-brane solution (231) has 16 Killing spinors, and only is invariant under half of the supersymmetry transformations. Solutions with residual supersymmetry are called BPS solutions, and solutions which preserve half of the supersymmetry are called '$\frac{1}{2}$ BPS solutions'.

The Bogomol'nyi bound (247) can be shown to follow from supersymmetry. In this context it is then also called the BPS bound. In theories where the supersymmetry algebra contains central charges, (247) is a relation between the mass or tension of a state and its central charge. In our case the charges carried under the R-R gauge fields are such central charges. The representations of the supersymmetry algebra fall into distinct classes, depending on whether they saturate the bound or not. Representations which saturate the bound are called short representations or BPS representations. Since BPS states have the minimal tension possible for their charge they are absolutely stable. This minimization of energy also accounts for the existence of static multicentered solutions.

So far we have restricted ourselves to p-brane solutions with $0 \leq p \leq 2$. There is a second class, where the solution (231) and the other formulae take the same form, but with p replaced by $\tilde{p}$ with $4 \leq \tilde{p} \leq 6$. The field strength $F_{\tilde{p}+2}$ in equation (231) is the $\star$–dual of $F_{p+2} = dA_{p+1}$. Since $F_{\tilde{p}+2} = \star F_{p+2}$ implies $p + \tilde{p} + 4 = D = 10$, each of the so-called electric solutions with $p = 0, 1, 2$ has a dual magnetic solution with $\tilde{p} = 6, 5, 4$.

There is also a solution with $p = 3$. The five-form gauge field is selfdual, and the solution for F_5 is from (231) by adding the $\star$-dual of the right-hand side of the equation. The solutions for the metric and for the dilaton are not modified. Note that for $p = 3$ the dilaton is constant. The three-brane solution is not singular at $r = 0$. Instead one has a regular horizon, and the geometry is asymptotic to $AdS^5 \times S^5$. This geometry has 32 Killing spinors and is fully supersymmetric. The interior of this geometry is isometric to the exterior, in particular it is non-singular. Since the field strength is selfdual, the three-brane carries an equal amount of electric and magnetic charge (it is not only dyonic, carrying both electric and magnetic charge, but selfdual).

Electric and magnetic charges are subject to a generalized Dirac quantization condition, which can be found by generalizing either the Dirac string or the Wu-Yang construction known from four-dimensional magnetic monopoles. In our conventions the condition is:

$$(2\pi)^7 g_S^2 (\alpha')^4 \, \hat{Q}_p \hat{Q}_{\tilde{p}} \in 2\pi \mathbf{Z} \ . \tag{253}$$

This fixes the possible magnetic charges in terms of the electric charges. Using T-duality and S-duality one can fix the electric and magnetic charge units. T-duality is a symmetry that can be proven to hold in string perturbation theory. It acts on our solutions by transforming p-branes into (p$\pm$1)-branes. In this way one can relate the tensions and charges of all R-R charged p-branes. S-duality is a conjectured non-perturbative symmetry of IIB string theory. It relates the R-R one-brane to a solution which describes the fundamental IIB string. This way one relates the fundamental unit of R-R one-brane charge to the charge carried by a fundamental IIB string under the NS-NS B-field. The resulting R-R p-brane charge units are given by μ_p (245) and satisfy Dirac quantization in a minimal way: $\mu_p \mu_{\tilde{p}} (2\pi)^7 (\alpha')^4 = 2\pi$. Thus N_p in (245) is an integer which counts multiples of the fundamental R-R charge. When using Q_p instead of $\hat{Q}_p$ to measure charges then c_p as defined in (235) is the unit charge.

We now summarize the R-R charged p-brane solutions of type II string theories:

Theory	R-R potential	electric sol.	magnetic sol.
IIA	$A_{(1)}$	$p = 0$	$p = 6$
IIB	$A_{(2)}$	$p = 1$	$p = 5$
IIA	$A_{(3)}$	$p = 2$	$p = 4$
IIB	$A_{(4)}$	$p = 3$ (selfdual)	

(254)

The R-R p-brane solutions have properties which qualify them as solitons: They are static, stable (BPS bound), regular (no naked singularities) solutions of the field equations and have finite tension. The three-brane has an additional property familiar from two-dimensional solitons: it interpolates between two vacua, Minkowski space at infinity and $AdS^5 \times S^5$ at the event horizon. (We call $AdS^5 \times S^5$ a vacuum, because it is maximally supersymmetric.) For solitons one expects that the tension depends on the coupling as $T \sim 1/g^2$. This is, for example, what one finds for monopoles in Yang-Mills-Higgs theories. In this respect the R-R p-branes show an unusal behaviour as their tension is proportional to the inverse coupling, $T_p \sim 1/g_S$, see (242). This behaviour is in between the one expected for a soliton $T \sim 1/g_S^2$ and the one of a fundamental string, $T \sim 1$, which is independent of the coupling.

One clue to this unexpected behaviour is that the fundamental coupling of three closed strings is – up to a constant – the square of the coupling of three open strings, see (141). Thus a R-R p-brane has the coupling dependence expected for a soliton in a theory of open strings. The type II string theories, as defined so far, are theories of oriented closed strings. Consider now an extension where one adds to the theory open strings with Dirichlet boundary conditions along p directions. If we manage to identify the corresponding D-p-branes with the R-R p-brane solutions, this provides a description of type II string theory in these solitonic backgrounds.

5.3 p-Branes and D-Branes

Surprising as it may be, the identification of R-R p-branes and D-branes can be supported by convincing arguments. Let us compare the known properties of these objects. R-R p-branes preserve half of supersymmetry and can be located at arbitrary positions in transverse space. The same is true for D-branes with p = $0, 2, 4, 6$ in type IIA and p = $1, 3, 5$ in type IIB string theory. The corresponding Killing spinors are constant and are given by (252). The translational symmetries trivially agree. These D-p-branes are BPS states and since the central charge associated with a BPS state with Killing spinors (252) is precisely the R-R charge, they must carry R-R charge. A crucial quantitative test is to compute the R-R charge carried by a single D-p-brane. To do so one has to compute the force due to exchange of R-R gauge fields between to D-p-branes.

One first computes an annulus diagram with Dirichlet boundary conditions on both boundaries. This diagram can be factorized in two ways: either as a sum over intermediate open strings, or as a sum over intermediate closed strings. In the closed string channel the diagram can be visualized as a cylinder (closed string propagator) ending on the two D-branes. In this picture it is obvious that one measures the total force between the D-branes resulting from the exchange of arbitrary closed string states. This amplitude vanishes, which tells us that the total force vanishes, as expected for a BPS state. To extract the long range part of the force one takes the two D-branes to be far apart and expands the amplitudes in the masses of the closed string states. Then the exchange of massless states dominates. In detail one finds an attractive force due to graviton and dilaton exchange which is cancelled exactly by a repulsive force due to exchange of rank (p + 1) tensor gauge fields. The static R-R forces correspond to a generalized Coulomb potential,

$$V_{\mathrm{R-R}} = \frac{Q_p}{r^{D-p-3}} = \frac{Q_p}{r^{7-p}} \ . \tag{255}$$

It turns out that one D-p-brane carries precisely one unit of R-R p-brane charge,

$$Q_p = c_p = \frac{(2\pi)^{7-p}}{(7-p)\omega_{8-p}} (\alpha')^{\frac{7-p}{2}} g_S \ . \tag{256}$$

This shows that one should identify a R-R p-brane of charge $N_p c_p$ with a system of N_p D-p-branes. People also have computed various other quantities, including the low energy scattering, absorption and emission (encoded in the so-called greybody factors) of various strings states on R-R p-branes and D-p-branes, and the low velocity interactions between p-branes and D-branes. All these tests have been successful.

Since p-branes are extended supergravity solutions with non-trivial space-time metric, whereas D-branes are defects in flat space-time, we should of course be more precise in what we mean by identification. We have seen that both kinds of objects have the same charges, tensions and low energy dynamics. They have the same space-time and supersymmetries and saturate the same BPS bound. Thus they seem to represent the same BPS state of the theory, but in different regions of the parameter space. A description in terms of N_p D-branes works within string perturbation theory. In presence of D-branes the effective string loop counting parameter is $N_p g_S$ instead of g_S. The reason is as follows: as we have seen in Sect. 3 each boundary component gives rise to a factor g_S in scattering amplitudes. In a background with D-branes every boundary component can end on each of the N_p D-branes and therefore g_S always occurs multiplied with N_p. Since we are interested in describing macroscopic objects with large N_p, we need to impose that $N_p g_S$ is small in order to apply perturbation theory.

Thus we are in the perturbative regime if

$$N_p g_S \ll 1 \ . \tag{257}$$

Using the Schwarzschild radius (239) we see that this equivalent to

$$r_S \ll \sqrt{\alpha'} \ , \tag{258}$$

which means that the gravitational scale is much smaller then the string scale. This explains why one does not see any backreaction of the D-branes on the space-time in string perturbation theory. D-branes have a finite tension and couple to gravity, but the deviation from flat space caused by backreaction is only seen at scales of the order r_S. The only length scale occuring in string perturbation theory is $\sqrt{\alpha'}$ and this is the minimal scale one can resolve when probing D-branes with strings.

The R-R p-branes are solutions of the type II effective actions. These are valid at string tree level and therefore we need to be in the perturbative regime (of the closed string sector), $g_S < 1$. Moreover we have neglected α'-corrections, which become relevant when the curvature, measured in string units, becomes large. The condition for having small curvature is

$$r_S \gg \sqrt{\alpha'} \ , \tag{259}$$

or, equivalently,

$$N_p g_S \gg 1 \ , \tag{260}$$

which is opposite to (257, 258). The p-brane solution is valid in the regime of the low energy effective field theory, where stringy effects can be neglected.

Between the two regimes one can interpolate by changing the string coupling g_S, while keeping the charge N_p fixed. In general it is not clear that one can believe in the results of such interpolations. But in our case we know that the p-brane/D-brane is the object of minimal tension for the given charge. As a BPS state it sits in a special BPS multiplet. There is no mechanism compatible with supersymmetry through which this state could decay or become a non-BPS state. Besides these arguments, various quantities have been computed in both regimes and agree with one another.

In string perturbation theory one also has D-branes with $p > 6$. Therefore one might wonder whether the corresponding objects also exist as p-branes.

The answer is yes, though these so-called large branes have somewhat different properties than the other branes. For example the seven- and eight-brane are not flat in the transverse dimensions. The reason is that there are no harmonic functions in transverse space that become constant at infinity (this is similar to black holes in $D < 4$). The seven-brane carries magnetic charge under the IIB R-R scalar A. Its electric partner is a (-1)-brane, the D-instanton. The eight-brane does not have a local source. It is a domain wall solution which separates regions where the IIA mass parameter (which is similar to a cosmological constant) takes different values. The nine-brane is flat space.

5.4 Further Reading

The type II effective actions and the corresponding p-brane solutions can be found in the book [3]. For extensive reviews of BPS-branes in supergravity and string theory, see [11,47,13,12].

6 Outlook

In this final section we give an outlook on more recent developments.

6.1 Eleven-Dimensional M-Theory

Besides the R-R charged p-branes, type II string theories contain various other BPS solutions. Since all these carry central charges of the supersymmetry algebra, they can be constructed systematically. The other string theories also have their BPS solitons. Combining perturbative string theory with the knowledge about the BPS states one can show that the strong coupling behaviour of any of the five string theories can be described consistently by a dual theory. Moreover, one can interrelate all five superstring theories by such string dualities. These dualities have not been fully proven yet, but one has compared various accessible quantities and all these tests have been successful. The dualities give a coherent picture where all perturbative string theories are limits of one single underlying theory.

This is by now a huge subject, which deserves a separate set of lectures. Here we will only illustrate it by reviewing Witten's analysis [18] of the strong coupling behaviour of type IIA string theory. Consider the spectrum of finite mass objects in IIA string theory. It starts with the massless IIA supergravity multiplet, then comes an infinite series of excited string states with masses (198),

$$\alpha' M^2 \sim N \;, \tag{261}$$

where $N = 1, 2, \ldots$. As further finite mass objects the theory contains states with N_0 D-0-branes, with masses (242),

$$\alpha' M^2 \sim \frac{N_0}{g_S} \;. \tag{262}$$

(One can show that there are no bound states at threshold, so the states with $N_0 > 1$ are N_0-particle states.) In the perturbative regime, $g_S \ll 1$, the D-0-branes are very heavy. But when extrapolating to strong coupling, $g_S \to \infty$, they become much lighter than any perturbative excitation. Since the D-0-branes are BPS-states, we know that the mass formula (262) is not modified at strong coupling. For very large g_S one gets a quasi-continuum of D-0-brane states above the massless supergravity multiplet. The collective modes of a D-0-brane sit in a so-called short multiplet of the IIA supersymmetry algebra. Short multiplets are special massive multiplets which saturate the BPS bound. They have less components than generic massive multiplets. The multiplet of the D-0-brane is a massive version of the supergravity multiplet: it has the same number of states and the same spin content. Thus the low energy, strong coupling spectrum looks like the Kaluza-Klein spectrum obtained by dimensional reduction of an eleven-dimensional theory. The only candidate is eleven-dimensional supergravity, the unique supersymmetric theory in eleven dimensions. When comparing the low energy, strong coupling spectrum of IIA string theory to the Kaluza-Klein spectrum of eleven-dimensional supergravity one finds that both agree, provided one relates the string coupling to the radius R_{11} of the additional space dimension according to,

$$g_S^2 = \left(\frac{R_{11}}{L_{\mathrm{Pl}}} \right)^3 \tag{263}$$

and the string scale α' to the eleven-dimensional Planck length L_{Pl} according to:

$$\alpha' = \frac{L_{\mathrm{Pl}}^3}{R_{11}} \, . \tag{264}$$

The eleven-dimensional Planck length is defined through the eleven-dimensional gravitational coupling by: $\kappa_{(11)}^2 = L_{\mathrm{Pl}}^9$. The relation between the eleven-dimensional metric and the IIA string frame metric is:

$$ds_{11}^2 = e^{2\Phi/3} \left(ds_{\mathrm{IIA,Str}}^2 + (dx^{11} - A_\mu dx^\mu)^2 \right) \, , \tag{265}$$

where Φ is the IIA dilaton and the Kaluza-Klein gauge field A_μ becomes the R-R one form.

This indicates that the strong coupling limit of IIA string theory is an eleven-dimensional theory, called M-theory. We do not have enough information to give a complete definition, but we know that M-theory has eleven-dimensional supergravity as its low energy limit. There must be additional degrees of freedom, because eleven-dimensional supergravity is not consistent as a quantum theory. Even without a complete definition of M-theory, one can find more evidence for the duality. Eleven-dimensional supergravity has BPS solitons, which properly reduce under dimensional reduction to the solitons of IIA string theory. In particular it has a supersymmetric membrane solution, called the M-2-brane, which reduces to the fundamental IIA string.

6.2 String Dualities

Let us now consider the other string theories. What about type IIB? The theory has maximal supersymmetry, and its massless spectrum cannot be obtained by dimensional reduction from a higher-dimensional supersymmetric theory. The only obvious possibility is that it is selfdual, which means that the strong and weak coupling limits take the same form. One can show that inverting the coupling, $g_S \to g_S^{-1}$, preserves the form of the action and is a symmetry of the BPS spectrum if one simultanously interchanges the fundamental IIB string with the D-1-brane. The strong coupling limit is again a IIB string theory, with solitonic strings (D-1-branes) now playing the role of the fundamental objects. The transformation relating weak and strong coupling is called S-duality and works the same way as the Montonen-Olive duality in four-dimensional $N = 4$ Super-Yang-Mills theory. It has also been verified that S-duality is respected by instanton corrections to string amplitudes.

In a similar way, the strong coupling limit of the type I string is the heterotic string with gauge group $SO(32)$, and vice versa. We already saw that both theories have the same massless spectra, while the perturbative massive spectra and interactions were different. Both theories cannot be selfdual (for example, inverting the string coupling does not preserve the form of the effective action). But once solitonic BPS states are included, the BPS spectra are equal and reversing the coupling relates the two effective actions. The heterotic $SO(32)$ string is identified with the D-1-brane of type I.

What is left is to determine the strong coupling behaviour of the $E_8 \times E_8$ heterotic string. This turns out to be again eleven-dimensional M-theory but this time compactified on an interval instead of a circle. The interval has two ten-dimensional boundaries, on which ten-dimensional vector multiplets with gauge group E_8 are located. This is also known as Horava-Witten theory.

Let us summarize the strong-coupling limits of the five supersymmetric string theories:

String theory	Strong coupling dual
IIA	M-theory on circle
IIB	IIB
I	Heterotic $SO(32)$
Heterotic $SO(32)$	I
Heterotic $E_8 \times E_8$	M-theory on intervall

(266)

These dualities fall into two classes: either one has a relation between strong and weak coupling. This is called S-duality. Or the coupling is mapped to a geometric datum, the radius of an additional dimension. There is a third type of string duality, which leads to further relations between string theories. It is called T-duality and relates weak coupling to weak coupling, while acting non-trivially on the geometry. Since weak coupling is preserved, one can check that T-duality is preserved in perturbation theory. By T-duality, the IIA string theory compactified on a circle of radius R is equivalent to IIB string theory compactified on a circle of inverse radius in string length units, $\tilde{R} = \frac{\alpha'}{R}$. One can take the decompactification limit and obtain ten-dimensional IIB theory as the zero radius limit of compactified IIA theory and vice versa. In the same way one can relate the two heterotic string theories. When acting on open strings, T-duality exchanges Neumann boundary conditions with Dirichlet boundary conditions. Therefore the T-dual of type I string theory is a theory containing open strings which are coupled to D-branes. Though one might consider this as a solitonic sector of type I theory, it is sometimes called type I' theory. Let us summarize the T-duals of the five supersymmetric string theories:

String theory	T-dual theory
IIA	IIB
IIB	IIA
I	I'
Heterotic $SO(32)$	Heterotic $E_8 \times E_8$
Heterotic $E_8 \times E_8$	Heterotic $SO(32)$

(267)

Finally there is yet another relation between IIB string theory and type I. Type IIB has supersymmetric D-9-branes. These D-branes are space-filling, they correspond to adding open strings with Neumann boundary conditions in all directions. From our earlier discussion we know that the only consistent coupling between open and closed superstrings is a non-oriented theory with gauge group $SO(32)$, namely type I. This can be realized as a configuration in IIB string theory, where one adds 32 D-9-branes together with additional non-dynamical objects, so-called orientifold planes, which reverse world-sheet parity. Type I string theory is an 'orientifold' of type IIB. More generally, after introducing D-branes and orientifolds, the type I, IIA and IIB string theories can be considered as one theory in different backgrounds, which can be transformed into another by T-duality and orientifolding. The type I' theory, which we introduced above as the T-dual of type I, can also be obtained as an orientifold of type IIA. Therefore type I' and type I theory are also called type IA and type IB.

Thus we see a bigger picture emerging once we include the BPS solitons of the five supersymmetric string theories. All theories are related to one another and to eleven-dimensional M-theory, and all strong couplings limits can be consistently described. Therefore one believes today that the different string theories are perturbative limits of one single underlying theory. Due to the role of D-branes and since there is an eleven-dimensional limit, which cannot be described by perturbative string theory, one prefers to call it M-theory.

6.3 Further Reading

String dualities and how they relate the five supersymmetric string theories to one another are discussed in the book [3] and in various lectures notes. The paper [48] gives a nice overview of the various dualities that we mentioned above. The lectures [46] approach the subject from the side of effective supergravity theories and string compactifications, whereas [47] is an introduction to supergravity which also covers branes and string dualities. Other lecture notes on string theory and string dualities are [49–51].

T-duality, which we only mentioned briefly in these lectures is reviewed at length in [16]. The role of combined T- and S-dualities, then called U-dualities, in string and M-theory compactifications is reviewed in [17]. D-branes and their applications are discussed in [52,53]. For a recent review of open strings, see [19]. The lectures [14] are devoted to the description of BPS black holes in string theory. They also cover the ten-dimensional brane solutions of type II string theories and how they are related by T- and S-duality. BPS solutions of eleven-dimensional supergravity (M-branes) and their relation to the brane solutions of type II string theory are explained in [12].

6.4 Lightning Review of Further TOPICS

Let us finally mention areas of active research together with some references, which might be useful for the interested reader.

What Is M-Theory? So far M-theory was characterized by its relation to various perturbative string theories and through its eleven-dimensional low energy limit, supergravity. The fundamental open question is how to define M-theory without recourse to a particular background, perturbation theory or particular limits. The recent developments show that besides strings also various branes have to be taken into account as dynamical objects. The question which remains open is which of these objects are truly fundamental. When considering all p-branes as equally fundamental as strings, one immediately faces the problems of how to quantize higher-dimensional objects. Among p-branes, strings ($p = 1$) and particles ($p = 0$) are singled out, because their world-volume theories are free as long as the background geometry is flat. This underlies the power of string perturbation theory. The situation is completely different for higher-dimensional branes ($p > 1$), where the world-volume theory is a complicated interacting theory, even in a flat background. Therefore no analogon of string perturbation theory for these objects has been developed so far. Alternatively, one particular kind of brane might be the fundamental object, whereas all others are obtained by dimensional reduction or as solitons. There are two candidates for which concrete proposals have been made: the supermembrane and the D-0-brane.

The Supermembrane. Eleven-dimensional supergravity has a solitonic two-brane solution, called the supermembrane or the M-2-brane. The three-dimensional action for the collective modes of this solution contains a Nambu-Goto term and Wess-Zumino term, which describe the coupling to gravity and to the three-form gauge field of eleven-dimensional supergravity. One can then try to treat this membrane as a fundamental object in an analogous way to the fundamental string in string theory. Moreover, one can get back the IIA string by dimensional reduction. Supermembrane theory is much more complicated than string theory, because the world-volume theory does not become free in a flat background, as discussed above. Also note that there is no local Weyl invariance for p-branes with $p \neq 1$. Therefore there is no conformal world volume action and no analogon of the Polyakov formulation.

At the WE-Heraeus-Seminar, supermembrane theory was the subject of lectures by Hermann Nicolai. A pedagogical introduction to the subject, which also covers the relation to other approaches to M-theory is provided by his Trieste lectures [55].

Matrix Theory. In the matrix-theory formulation of M-theory, also called M(atrix) theory, the D-0-brane is the fundamental object. More precisely, there is a conjecture due to Banks, Fischler, Shenker and Susskind [22], which claims that eleven-dimensional M-theory in the infinite momentum frame is given exactly by the limit $N \to \infty$ of the supersymmetric $U(N)$ quantum mechanics describing a system of N D-0-branes.

M(atrix) theory can be viewed as an alternative formulation of supermembrane theory, since the finite–N–M(atrix) model Hamiltonian is an approximation of the supermembrane Hamiltonian. In M(atrix) theory multi-membrane states are described by clusters of D-0-branes. Conversely, D-0-branes are contained in supermembrane theory as Kaluza-Klein modes of the eleven-dimensional supergravity multiplet, which consists of the zero-mass states of the supermembrane. Besides [55], lectures on M(atrix) theory are [56–58].

Black Holes. While the fundamental definition of M-theory remains to be found, string theory and D-branes have been applied to a variety of problems in gravity, field theory and particle physics. One of the most prominent applications is the description of black holes through D-branes, which elaborates on the relation between D-branes and p-brane solutions discussed in Sect. 5. Starting from p-branes in ten dimensions one can obtain four-dimensional black holes by dimensional reduction. Performing the same reduction with the corresponding D-brane configuration, one gets a description of the system where the microscopic degrees of freedom are known. This can be used to compute the entropy of the black hole: one counts the number N of microstates, i.e., excitations of the system, which belong to the same macrostate, i.e., the same total energy, charge and angular momementum:

$$S = \log N \ . \tag{268}$$

In practice the statistical entropy N is evaluated asymptotically for very large black hole mass.

The result can be compared to the Bekenstein-Hawking entropy of the black hole, which is given in terms of the area A of the event horizon,

$$S_{\mathrm{BH}} = \frac{A}{4} \,. \tag{269}$$

One finds that the two entropies agree, $S = S_{\mathrm{BH}}$, which confirms that the D-brane picture correctly captures the microscopic degrees of freedom of the black hole [23]. As mentioned above S is evaluated asymptotically, but we would like to stress that the resulting S matches exactly with the Bekenstein-Hawking entropy. This is in contrast to other approaches, where both entropies have the same dependence on parameters, while the numerical prefactor of the statistical cannot be determined precisely.

One can also compute and compare sub-leading contributions to both entropies. Corrections to the statistical entropy have been computed for Calabi-Yau compactifications of type II string theory and eleven-dimensional M-theory [25,26], see also [27]. These match precisely with corrections to the macroscopic black hole entropy, which are due to higher curvature terms in the effective action [28,29]. These higher curvature terms modify the entropy in two ways. The first is an explicit modification of the black hole solution and, hence, of the area A. The second is a modification of the area law (269). As pointed out by R. Wald [24], the validity of the first law of black hole mechanics in presence of higher curvature terms requires a modified definition of black hole entropy. (The first law of black hole mechanics formulates the conservation of energy. It expresses adiabatic changes of the mass to changes in terms of parameters of the black hole solution.) Both effects, the explicit change of the solution and the modified definition of the entropy, change the entropy in a complicated way, but the combined correction is relatively simple and precisely matches the statistical entropy. This is reviewed in [15].

Besides entropy, the D-brane picture has been used to compute Hawking radiation and greybody factors (see [61,44] for review and references). This is possible for branes which are close to the BPS limit. In the D-brane picture one can compute the emission, absorption and scattering of closed string states by a D-brane. Again one finds agreement with a semiclassical treatment of the corresponding black hole solutions. Note, however, that the method only applies to D-branes and p-branes which are close to the BPS limit. The generalization to generic black holes remains an open problem, though various proposals have been made. One idea, which applies to black holes without R-R charge, is a correspondence principle between black holes and fundamental strings [30,31]. The idea is that a black hole evaporates through Hawking radiation until its size reaches the string scale where it converts into a highly excited fundamental string. This is supported by the observation that the entropies of black holes and fundamental strings of equal mass match precisely when the Schwarzschild radius equals the string length. Another idea [32] is to use string dualities to map four-(and five-)dimensional black holes to three-dimensional black holes (BTZ black holes [33]). Three-dimensional gravity does not have local degrees of freedom, because the action is a total derivative. If the space-time has boundaries one gets boundary degrees of freedom which can be described by a two-dimensional conformal field theory. Treating the horizon as a boundary, this can be used to compute the statistical entropy of three-dimensional black holes [34]. The dualities that one needs to connect these three-dimensional to four-dimensional black holes are slightly more general then those mentioned so far. In particular they change the asymptotic geometry of space-time, so that one can map a higher-dimensional black hole to a lower-dimensional one (times an internal, compact space). One can argue that these transformations do not change the thermodynamic properties. Moreover one finds explicitly that the Bekenstein-Hawking entropy of the four-dimensional Schwarzschild black hole is matched by the state counting of the dual three-dimensional black hole. A related approach is to use dualities to map Schwarzschild black holes to brane configurations [35]. Finally, the microscopic entropy of Schwarzschild black holes has also been computed using Matrix theory, see [58] for review and references.

The most general and most promising approach to generic black holes is the AdS-CFT correspondence [38–40]. This correspondence and its generalizations relate D-dimensional gravitational backgrounds to $(D-1)$–dimensional field theories. One of the roots of this idea is the so-called holographic principle [36,37], which claims that the physics beyond the horizon of a black hole can be described in terms of a field theory associated with its horizon. The D-brane picture of black holes can be viewed as a realization of this idea, because here the interior region of the black hole has disappeared, while interactions of the exterior region with the black hole are described as interactions between closed strings in the bulk with open strings on the brane. A more general version of the holographic principle is that gravity can always be described in terms of a lower-dimensional field theory. The AdS-CFT correspondence, which we briefly describe below, can be viewed as an attempt to realize this idea.

More about black holes in string theory can be found in [14] and in other reviews of the topic including [59–62,44,63,15,58] and Sect. 14.8 of [3].

The AdS-CFT Correspondence and Its Generalizations. The AdS-CFT correspondence is another consequence of the relation between D-branes and p-brane solutions. Its most simple version is obtained by considering a system of N D-3-branes and taking the limit $\alpha' \to 0$, while Ng_S and R/α' are kept fixed. Here g_S is the string coupling and R the characteristic scale of separation between the branes. In the D-brane picture gravity and massive string excitations decouple and one is left with the effective theory of the massless open string modes, which is a four-dimensional $\mathcal{N} = 4$ supersymmetric $U(N)$ gauge theory in the large N limit. The corresponding limit in the p-brane regime is the near horizon limit, where the geometry takes the form $AdS^5 \times S^5$. The low energy excitations are described by supergravity on AdS^5. This observation motivated Maldacena's conjecture [38]: five-dimensional supergravity on AdS^5 is a dual description of four-dimensional $\mathcal{N} = 4$ supersymmetric $U(N)$ gauge theory, the latter being a conformal field theory. AdS^5 has an asymptotic region which can be identified with (the conformal compactification of) four-dimensional Minkowski space.

This is called the boundary, and the conformal field theory is located there. One finds a correspondence between fields $\phi(x_{(5)})$ of the bulk supergravity theory and operators $\mathcal{O}(x_{(4)})$ of the Yang-Mills theory on the boundary. (Here $x_{(5)}$ are coordinates on the five-dimensional bulk and $x_{(4)}$ are coordinates on the four-dimensional boundary.) A quantitative version of the conjecture, due to Gubser, Klebanov, Polyakov [39] and Witten [40], states that the generating functional for the correlators of operators $\mathcal{O}(x_{(4)})$ with sources $\phi_0(x_{(4)})$ is given by the partition function of the supergravity theory, evaluated in the background $\phi(x_{(5)})$ with boundary values $\phi(x_5)|_{\text{Boundary}} = \phi_0(x_{(4)})$, according to:

$$\left\langle e^{\int d^4x \phi_0(x_{(4)}) \mathcal{O}(x_{(4)})} \right\rangle = Z\left(\phi(x_{(5)})\right) . \tag{270}$$

There are various generalizations of this basic form of the correspondence, which relate other gravitational backgrounds to other gauge theories. One particular extension of the AdS-CFT correspondence relates five-dimensional domain wall geometries to renormalization group flows in non-conformal gauge theories. In this setup the coordinate transverse to the domain wall corresponds to the energy scale of the gauge theory [41,42]. More recently, maximally supersymmetric gravitational wave backgrounds have moved to the center of interest [43].

Extensive reviews of the AdS-CFT correspondence can be found in [44] and [45].

Brane Worlds. D-branes provide a new option for model building in particle physics. One can localize some or all matter and gauge fields of the standard model on a three-brane, while gravity propagates in the higher-dimensional bulk. Such models have the interesting feature that the size of the extra dimensions can be quite large, even in the sub-mm range. Moreover one can have a fundamental (higher-dimensional) Planck scale of 1 TeV, which provides a new approach to the gauge hierarchy problem. A low gravitational scale of 1 TeV leads to spectacular predictions, like the mass production of black holes at the LHC. Therefore brane worlds have been a main activity in the string and particle physics community over the last years. One should stress here that though TeV-scale gravity is possible within string theory, it is not predicted.

There is a huge variety of brane world models, which range from phenomenological models to models with explicit realization in string or M-theory, see for example [64–69]. In one variant, the so-called Randall-Sundrum model (RS II model [68], to be precise), the extra dimensions are curved in such a way that gravity is confined on the brane in a similar way as matter fields. This opens the possibility of extra dimensions which are arbitrarily large, though invisible at low energies.

At the WE-Heraeus Seminar brane worlds were the subject of the lectures given by I. Antoniadis and A. Barvinsky, while J. Gundlach reviewed tests of Newton's law at short distances. A nice review of mass scales and the possible sizes of extra dimensions in string theory can be found in [20]. Experimental signatures of large extra dimensions are discussed in [21]. One particular type of brane worlds, which occur in Calabi-Yau compactifications of Horava-Witten theory, is reviewed in [70]. The lectures [72] give an introduction to brane worlds and warped compactifications.

Compactifications and Phenomenology. D-branes and p-branes have considerably extended the framework of string compactifications, which aim to explain how our four-dimensional world is embedded into the fundamental ten- or eleven-dimensional theory. Whereas ten years ago string phenomenology was synonymous with the study of the heterotic $E_8 \times E_8$ string, compactified on complex three-dimensional Calabi-Yau manifolds, one now has various other options to consider. Besides brane worlds one can study compactifications where part of the standard model particles are not string modes but descend from p-branes wrapped on internal p-cycles. Switching on background fluxes of antisymmetric tensor fields, one obtains warped compactifications, where the characteristic length scale of four-dimensional space-time becomes dependent on the position in the internal space. A particular class of non-perturbative IIB backgrounds can be described purely geometrically in terms of so-called F-Theory.

The central problem of string compactifications is still the problem of vacuum degeneracy. As we have seen, the vacuum expectation value of the dilaton is not fixed at string tree level. In supersymmetric theories this holds to all orders in perturbation theory. Similarly, string compactifications in general have several scalar fields, called moduli, which parametrize the shape and size of the internal manifold and enter into the couplings of the effective field theory. The vacuum expectation values of these fields are not fixed, as long as supersymmetry is unbroken. This ruins the predictive power that the theory has in principle, and leads to continuous families of degenerate vacua. Once supersymmetry is broken the moduli get fixed, but there is a number of issues to be addressed: one needs to understand the dynamical mechanism behind supersymmetry breaking, which requires to understand the theory non-perturbatively. The potential generated for the dilaton and for the moduli should have stable vacua and no runaway behaviour. One needs sufficiently large masses or sufficiently small couplings for the moduli to avoid contradiction with empirical data. Moreover, in string theory supersymmetry is closely related to the absence of the tachyon, which one does not want to reintroduce. One also wants that supersymmetry breaking occurs at a specific scale, the most popular scenario being low energy supersymmetry where the supersymmetric partners have masses of about 1 TeV. D-branes, p-branes and other new developments have added a variety of new ways to address these problems, but a definite solution remains to be found.

String compactifications on Calabi-Yau manifolds are reviewed in [71]. Lectures on warped compactifications and brane worlds can be found in [72]. F-theory is for example explained in [49]. For an introduction to string and M-theory particle phenomenology, see for example [73,74].

Geometric and D-Brane Engineering, D-Branes, and Non-commutative Field Theory. In addition to the AdS-CFT correspondence, string theory has led to other new approaches to gauge theories and other field theories. In geometric engineering [77] one starts from branes wrapped on cycles in an internal space, which typically is a Calabi-Yau manifold, whereas in D-brane engineering [78] one studies D-brane configurations in a non-compact space-time. In both cases one takes a low energy limit (similar to the one discussed above in the context of the AdS-CFT correspondence) to decouple gravity.

Another direction stimulated by string theory and D-branes is gauge theory on non-commutative space-times. As mentioned in the lectures, the effective action for a D-brane is of Born-Infeld type. It has been argued that this can be reformulated as a Yang-Mills theory on a non-commutative world volume, with a deformation parameter which is determined by the bulk $B_{\mu\nu}$ field of the closed string sector [79].

Geometric engineering is reviewed in [75,76], while gauge theory on non-commutative space-times is reviewed in [80]. For extensive lectures on D-branes, see [52,53].

Cosmology. Whereas string compactifications usually aim at finding four-dimensional Minkowski space with a realistic particle spectrum from string theory, one should of course try to do better. Cosmological solutions of string theory should shed light on the issue of the initial singularity, describe an inflationary phase (or an alternative mechanism which takes care of the problems of the old hot big bang model), further describe the post-inflationary phase and explain the smallness of the cosmological constant. These problems have been mostly neglected by string theorists for a long time, but nowadays they find increasing interest, due to both new cosmological data and new theoretical developments. In particular branes have been invoked for either providing the mechanism for inflation or for providing an alternative to inflation.

Reviews of string cosmology can be found in [81–83].

The Challenge from de Sitter Space. Since there is empirical evidence in favour of a small, positive cosmological constant, there has been a considerable interest in string theory in de Sitter space over the last few years. De Sitter space is a challenge for several reasons. First, most successful applications of string theory to gravity depend on supersymmetry, but supersymmetry is completely broken in presence of a positive cosmological constant. Second, de Sitter space has cosmological horizons, and the perturbative formalism which works for Minkowski space as explained in Sect. 3 cannot be applied. Therefore de Sitter space requires a significant step beyond that framework. For a review see [54].

Tachyon Condensation and String Field Theory. As observed several times in these lectures, the appearance of tachyons is a generic feature of string theories when there is no supersymmetry. Since the mass squared of a scalar particle is given by the curvature of its potential at the stationary point one is expanding around, this shows that one tries to expand the theory around a local maximum of the potential. Depending on the global form of the potential, the theory might be unstable, or it might be that the scalar field rolls to a minimum. This is referred to as tachyon condensation.

Tachyons do not only occur in the ground state of bosonic string theories, but also in D-branes configurations which are not BPS states (non-BPS D-branes systems are reviewed in [84]). Work starting with a paper by A. Sen [85] provided strong evidence that tachyon condensation occurs in unstable non-BPS configurations of D-branes. Such systems have tachyonic open string states which condense. The resulting stable vacuum is the closed string vacuum, whereas the D-branes have decayed and therefore open strings are absent. This work makes use of string field theory, which for a long time was mostly neglected, because it is very complicated and was believed of little practical use. The renewed interest in string field theory might bring us one step forward towards a non-perturbative and background-independent formulation of M-theory.

Acknowledgements

First of all I would like to thank the organisers of the 271-th WE-Heraeus seminar for organizing this very lively and stimulating school. I would also like to thank the participants for their numerous questions and remarks, which I have tried to take into account when writing these lecture notes. The same applies to all the participants of my string theory course at the University of Jena. My special thanks goes to Frank Saueressig and Christoph Mayer for their very helpful comments on the manuscript.

References

1. M.B. Green, J.H. Schwarz and E. Witten: *Superstring Theory* (Cambridge University Press, Cambridge 1984)
2. D. Lüst and S. Theisen: *Lectures on String Theory* (Springer, Berlin Heidelberg 1989)
3. J. Polchinski: *String Theory* (Cambridge University Press, Cambridge 1998)
4. M. Kaku: *Introduction to Superstrings and M-Theory* (Springer, Berlin Heidelberg 2000)
5. M. Kaku: *Conformal Fields and M-Theory* (Springer, Berlin Heidelberg 2000)
6. P. Ginsparg: 'Applied Conformal Field Theory'. In: *Fields, Strings and Critical Phenomena, Les Houches Summer School in Theoretical Physics, Session 49, Les Houches, France, June 28–August 5, 1988,* ed. by E. Brezin and J. Zinn-Justin (North-Holland, Amsterdam 1990)
7. E. D'Hoker and D.H. Phong: Rev. Mod. Phys. **60**, 917 (1988)
8. S. Weinberg: 'Covariant Path-Integral Approach to String Theory'. In: *Strings and Superstrings, 3rd Jerusalem Winter School for Theoretical Physics, Jerusalem, Israel, December 30, 1985–January 9, 1986*, ed. by S. Weinberg and T. Piran (World Scientific, Singapore 1988)
9. S. Albeverio, J. Jost, S. Paycha and S. Scarlatti: *A Mathematical Introduction to String Theory* (Cambridge University Press, Cambridge 1997)

10. M. Schlichenmaier: *An Introduction to Riemann Surfaces,Algebraic Curves and Moduli Spaces* (Springer, Berlin Heidelberg 1989)
11. K.S. Stelle: 'BPS Branes in Supergravity'. In: *High Energy Physics and Cosmology 1997, ICTP Trieste Workshop, Trieste, Italy, June 2–July 11, 1997.* ed. by E. Gava, A. Masiero, K.S. Narain, S. Randjbar-Daemi, G. Senjanovic, A. Smirnov, Q. Shafi. (World Scientific, Singapore 1998). Preprint: hep-th/9803116
12. P.K. Townsend: 'M-Theory from its Superalgebra'. In: *Strings, Branes and Dualities, NATO Advanced Study Institute, Cargese, France, May 26–June 14, 1997,* ed. by L. Beaulieu, P. DiFrancesco, M. Douglas, V. Kazakov, M. Picco, P. Windey. (Kluwer, Dordrecht 1999). Preprint: hep-th/9712004
13. M.J. Duff: 'Supermembranes'. Preprint: hep-th/9611203
14. T. Mohaupt: Class. Quant. Grav. **17**, 3429 (2000)
15. T. Mohaupt: Fortschr. Phys. **49**, 1 (2001)
16. A. Giveon, M. Porrati and E. Rabinovici: Phys. Rept. **244** 44 (1994). Preprint: hep-th/9401139
17. N.A. Obers and B. Pioline: Phys. Rept. **318**, 113 (1999). Preprint: hep-th/9809039
18. E. Witten: Nucl. Phys. B **433** 85 (1995). Preprint: hep-th/9503124
19. C. Angelantonj and A Sagnotti: 'Open Strings', Preprint: hep-th/0204089
20. I. Antoniadis: 'Mass Scales in String and M-Theory'. In: *Superstrings and Related Matters, ICTP Trieste Workshop, Trieste, Italy, March 22–30, 1999,* ed. by B. Green, J. Louis, K.S. Narain, S. Randjbar-Daemi (World Scientific, Singapore 2000)
21. I. Antoniadis and K. Benakli: Int. J. Mod. Phys. A **15**, 4237 (2000). Preprint: hep-ph/0007226
22. T. Banks, W. Fischler, S.H. Shenker and L. Susskind: Phys. Rev. D **55**, 5112 (1997) hep-th/9610043
23. A. Strominger and C. Vafa: Phys. Lett. B **379**, 99 (1996)
24. R.M. Wald: Phys. Rev. D **48**, 3427 (1993)
25. J. Maldacena, A. Strominger and E. Witten: JHEP **9712**, 002 (1997)
26. C. Vafa: Adv. Theor. Math. Phys. **2**, 207 (1998)
27. G.L. Cardoso, B. de Wit and T. Mohaupt: Nucl. Phys. B **567**, 87 (2000)
28. G.L. Cardoso, B. de Wit and T. Mohaupt: Phys. Lett. B **451**, 309 (1999)
29. G.L. Cardoso, B. de Wit, J. Käppeli and T. Mohaupt: JHEP **0012**, 019 (2000)
30. L. Susskind: 'Some Speculations about Black Hole Entropy in String Theory'. Preprint: hep-th/9309145
31. G.T. Horowitz and J. Polchinski: Phys. Rev. D **55**, 6189 (1997)
32. K. Sfetsos and K. Skenderis: Nucl. Phys. B **517**, 179 (1998)
33. M. Banados, C. Teitelboim and J. Zanelli: Phys. Rev. Lett. **69**, 1849 (1992)
34. S. Carlip: Phys. Rev. D **51**, 632 (1995), Phys. Rev. D **55**, 878 (1997)
35. R. Argurio, F. Englert and L. Houart: Phys. Lett. B **426**, 275 (1998)
36. G. 't Hooft: 'Dimensional Reduction in Quantum Gravity'. Preprint: gr-qc/9310026
37. L. Susskind: J. Math. Phys. **36**, 6377 (1995)
38. J. Maldacena: Adv. Theor. Math. Phys. **2**, 231 (1998). Preprint: hep-th/9711200
39. S.S. Gubser, I.R. Klebanov and A.M. Polyakov: Phys. Lett. B **248**, 105 (1998). Preprint: hep-th/9802109
40. E. Witten: Adv. Math. Phys. **2**, 253 (1998). Preprint: hep-th/9802150
41. H. Itzhaki, J.M. Maldacena, J. Sonnenschein and S. Yankielowicz: Phys. Lett. B **432**, 298 (1998). Preprint: hep-th/9802042
42. H.J. Boonstra, K. Skenderis and P.K. Townsend: JHEP **9901**, 003 (1999). Preprint: hep-th/9807137

43. D. Berenstein, J.M. Maldacena and H. Nastase: JHEP **0204**, 013 (2002). Preprint: hep-th/0202021
44. O. Aharony, S.G. Gubser, J. Maldacena, H. Ooguri and Y. Oz: Phys. Rept. **323**, 183 (2000)
45. E. D'Hoker and D.Z. Freedman: 'Supersymmetric Gauge Theories and the AdS/CFT Correspondence'. Preprint: hep-th/0201253
46. B. de Wit and J. Louis: 'Supersymmetry and Dualities in Various Dimensions'. In: *Strings, Branes and Dualities, NATO Advanced Study Institute, Cargese, France, May 26–June 14, 1997,* ed. by L. Baulieu, P. Di Francesco, M. Douglas, V. Kazakov, M. Picco, P. Windey. (Kluwer, Dordrecht 1999)
47. P.C. West: 'Supergravity, Brane Dynamics and String Duality'. In: *Nonperturbative Aspects of Strings, Branes and Supersymmetry, ICTP Conference on Super Five Brane Physics in* $5+1$ *Dimensions, Trieste, Italy, April 1–3, 1998,* ed. by M. Duff, E. Sezgin, C.N. Pope, B. Greene, J. Louis, K.S. Narain, S. Randjbar-Daemi, G. Thompson (World Scientific, Singapore 1999)
48. M. Haack, B. Körs and D. Lüst: 'Recent Developments in String Theory: From Perturbative Dualities to M-Theory'. Lectures given at the 4th National Summer School for Graduate Students of Theoretical Physics, Saalburg, Germany, August 31–September 11, 1998. Preprint: hep-th/9904033
49. A. Sen: 'An Introduction to Non-Perturbative String Theory'. In:*Duality and Supersymmetric Theories, Newton Institute Euroconference, Cambridge, England, April 7–18, 1997,* ed. by D.I. Olive, P.C. West (Cambridge University Press, Cambridge 1999). Preprint: hep-th/9802051
50. E. Kiritsis: 'Supersymmetry and Duality in Field Theory and String Theory'. In: *Particle Physics: Ideas and Recent Developments, NATO Advanced Study Institute, Cargese, France, July 26–August 7, 1999,* ed. by J.-J. Aubert, R. Gastmans, J.-M. Gaerard, (Kluwer, Dordrecht 2000). Preprint: hep-ph/9911525.
    E. Kiritsis: 'Introduction to Superstring Theory'. In: *Leuven Notes in Mathematical and Theoretical Physics. B9,* (University of Leuven Press, Leuven 1998). Preprint: hep-th/9709062.
    E. Kiritsis: 'Introduction to Non-Perturbative String Theory'. In: *Trends on Theoretical Physics, CERN–Santiago De Compostela–La Plata Meeting on Trends in Theoretical Physics, La Plata, Argentina, April 28–May 6, 1997*, ed. by H. Falomir, R.E. Gomboa Saravi, F.A. Schaposnik, (Amer. Inst. Phys. Woodbury, 1998). Preprint: hep-th/9708130
51. R.J. Szabo: 'Busstepp Lectures on String Theory'. Preprint: hep-th/0207142
52. C.P. Bachas: 'Lectures on D-Branes'. In: *Proceedings of the Newton Euroconference on Duality and Supersymmetric Theories, April 7–18, 1997, Cambridge, England,* ed. by D.I. Olive, P.C. West (Cambridge University Press, Cambridge 1999)
53. C.V. Johnson: 'D-Brane Primer'. Preprint: hep-th/0007170
54. M. Spradlin, A. Strominger and A. Volovich: 'Les Houches Lectures on De Sitter Space'. Preprint: hep-th/0110087
55. H. Nicolai and R. Helling: 'Supermembranes and M(atrix) Theory'. In: *ICTP Spring School on Nonperturbative Aspects of String Theory and Supersymmetric Gauge Theories, Trieste, Italy, March 23–31, 1998,* ed. by M. Duff, E. Sezgin, C.N. Pope, B. Greene, J. Louis, K.S. Narain, S. Randjbar-Daemi, G. Thompson (World Scientific, Singapore 1999). Preprint: hep-th/9809103
56. A. Bilal: 'M(atrix) Theory: a Paedagogical Introduction'. In: *Meeting on Quantum Aspects of Gauge Theories, Supersymmetry and Unification, Neuchatel, Switzerland, September 18–23, 1997,* ed. by J.-P. Derendinger, C. Lucchesi, (Wiley-VHC, Weinheim 1997) (Fortschr. Phys. **47** 1 (1999)). Preprint: hep-th/9710136

57. T. Banks: Nucl. Phys. Proc. Suppl. **67** 180 (1998). Preprint: hep-th/9710231
58. D. Bigatti and L. Susskind: 'Review of Matrix Theory'. In: *Strings, Branes and Dualities, NATO Advanced Study Institute, Cargese, France, May 26–June 14, 1997,* ed. by L. Baulieu, P. Di Francesco, M. Douglas, V. Kazakov, M. Picco, P. Windey. (Kluwer, Dordrecht 1999). Preprint: hep-th/9712072
59. R. D'Auria and P. Fré: 'BPS Black Holes in Supergravity'. Preprint: hep-th/9812160
60. D. Youm: Phys. Rept. **316**, 1 (1999). Preprint: hep-th/9710046
61. J. Maldacena: Black Holes in String Theory. PhD Thesis, Princeton University, Princeton (1996). Preprint: hep-th/9607235
62. G. Mandal: 'A Review of the D1/D5 System and Five Dimensional Black Hole from Supergravity and Brane Viewpoints'. Preprint: hep-th/0002184
63. K. Skenderis: 'Black Holes and Branes in String Theory'. Preprint: hep-th/9901050
64. V. Rubakov and M. Shaposhnikov: Phys. Lett. B **125**, 136 (1983)
65. N. Arkani-Hamed, S. Dimopoulos and G. Dvali: Phys. Rev. D **59**, 086004 (1999)
66. I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali: Phys. Lett. B **436**, 257 (1998)
67. L. Randall and R. Sundrum: Phys. Rev. Lett. **83**, 3370 (1999)
68. L. Randall and R. Sundrum: Phys. Rev. Lett. **83**, 4690 (1999)
69. A. Lukas, B.A. Ovrut, K.S. Stelle and S. Waldram: Phys. Rev. D **59**, 086001 (1999), Nucl. Phys. B **552**, 246 (1999)
70. B. Ovrut: '$N = 1$ Supersymmetric Vacua in Heterotic M-Theory'. Preprint: hep-th/9915115
71. B.R. Greene: 'String Theory on Calabi-Yau Manifolds'. In: *Fields, Strings and Duality, TASI 1996, Boulder, USA, June 2–28, 1996,* ed. by C. Efthimiou, B. Greene (World Scientific, Singapore 1997). Preprint: hep-th/9702155
72. S. Kachru: 'Lectures on Warped Compactifications and Stringy Brane Constructions'. In: *Strings, Branes, and Gravity, TASI 1999, Boulder, USA, May 31–June 25, 1999,* ed. by J. Harvey, S. Kachru, E. Silverstein, R. Edge, (World Scientific, Singapore 2001). Preprint: hep-th/0009247
73. J. Louis: 'Phenomenological Aspects of String Theory'. In: *Nonperturbative Aspects of String Theory and Supersymmetric Gauge Theories, ICTP Spring School, Trieste, Italy, March 23–31, 1998,* ed. by M. Duff, E. Sezgin, C.N. Pope, B. Greene, J. Louis, K.S. Narain, S. Randjbar-Daeme, G. Thompson (World Scientific, Singapore 1999)
74. M. Dine: 'TASI Lectures on M Theory Phenomenology. In: *String, Branes, and Gravity, TASI 99, Boulder, USA, May 31–June 25, 1999,* ed. by J. Harvey, S. Kachru, E. Silverstein, R. Egde (World Scientific, Singapore 2001). Preprint: hep-th/0003175
75. A. Klemm: 'On the Geometry behind $N = 2$ Supersymmetric Effective Actions in Four Dimensions'. In: *Duality – Strings and Fields, 33rd Karpacz Winter School of Theoretical Physics, Karpacz, Poland, February 13–22, 1997,* ed. by Z. Hasiewicz, Z. Jaskolski, J. Sobczyk (North-Holland, Amsterdam 1998). Preprint: hep-th/9705131
76. P. Mayr: 'Geometric Construction of $N = 2$ Gauge Theories'. In: *Meeting on Quantum Aspects of Gauge Theories, Supersymmetry and Unification, Neuchatel, Switzerland, September 18–23, 1997,* ed. by J.-P. Derendinger, C. Lucchesi, (Wiley-VHC, Weinheim 1997) (Fortschr. Phys. **47** 1 (1999)). Preprint: hep-th/9807096
77. A. Klemm, W. Lerche, P. Mayr, C. Vafa and N. Warner: Nucl. Phys. B477, 746 (1996). Preprint: hep-th/9604034;

S. Katz, A. Klemm and C. Vafa: Nucl. Phys. B **497**, 746 (1997). Preprint: hep-th/9609239;
S. Katz, P. Mayr and C. Vafa: Adv. Theor. Math. Phys. **1**, 53 (1998). Preprint: hep-th/9706110
78. A. Hanany and E. Witten: Nucl. Phys. B **492**, 152 (1997). Preprint: hep-th/9611230
79. N. Seiberg and E. Witten: JHEP **09**, 032 (1999)
80. M.R. Douglas and N.A. Nekrasov: Rev. Mod. Phys. **73** 977 (2002). Preprint: hep-th/0106048
81. J.E. Lidsey, D. Wands and E.J. Copeland: Phys. Rept. **337** 343 (2000). Preprint: hep-th/9909061
82. D.A. Easson: Int. Jour. Mod. Phys. A **16** 4803 (2001). Preprint: hep-th/0003086
83. M. Gasperini and G. Veneziano: 'The Pre-Big Bang Scenario in String Cosmology'. Preprint: hep-th/0207130
84. J.H. Schwarz: 'TASI Lectures on Non-BPS D-Branes Systemes'. In: *Strings, Branes, and Gravity, TASI 99, Boulder, USA, May 31–June 25, 1999,* ed. by J. Harvey, S. Kachru, E. Silverstein, R. Edge (Worlscientific, Singapore 2001). Preprint: hep-th/9908144
85. A. Sen: 'Tachyon Condensation on the Brane Antibrane System', JHEP **9809**, 023 (1998). Preprint: hep-th/9805170

Quantum Theory of Gravitational Collapse (Lecture Notes on Quantum Conchology)

Petr Hájíček

Institute for Theoretical Physics, University of Bern, Sidlerstrasse 5,
3012 Bern, Switzerland

Abstract. These notes consist of three parts. The first one contains the review of previous work on a gauge-invariant Hamiltonian dynamics of generally covariant models. The method is based on the exclusive use of gauge-invariant variables, the so-called Dirac observables, and on privileged dynamical symmetries such as the asymptotic time translation. The second part applies the method to the model of spherically symmetric thin shell of light-like substance in its own gravitational field following a paper by C. Kiefer and myself. A natural set of Dirac observables is chosen and the Hamiltonian defined by the time translation symmetry is calculated. In the third part, my construction of a version of quantum mechanics for the model is reviewed. The quantum evolution is unitary in spite of the classical theory containing black and white holes and singularities. The wave packet describing the quantum shell contracts, bounces and reexpands. The state of the quantum horizon is a linear combination of the "white" and "black" (past and future) apparent horizons.

1 Introduction

The issue of gravitational collapse includes not only the theory of black holes and of their classical and quantum properties, but also the serious problem of singularities. We believe that the singularities can be cured by quantum theory, but not by any semiclassical approximation thereof. The classical solutions near the singularity have to be changed strongly. New ideas concerning singularities are likely to influence the theory of black holes, too.

The rejection of WKB approximations forces us to use another kind of approximative schema. Simplified models constitute a promising alternative. Our method of dealing with the problem will, therefore, employ simplified models and a kind of effective theory of gravity. It does not worry about the final form of a full-fledged theory of quantum gravity. This need not be a completely unreasonable approach. Even if the ultimate quantum gravity theory were known, most calculations would still be performed by an approximation schema within some effective theory and for simplified models, such as the method of effective quantum theory in QCD. Another example of the method of simplified models used successfully today is the theory of atomic spectra which belongs, strictly speaking, to the field of quantum electrodynamics. Still, an infinite number of degrees of freedom can be safely frozen and the Schrödinger equation applied. Such methods can give useful hints for the gravitational collapse also because of the fact that the mainstay of the black hole geometry is formed by the purely

P. Hájíček, Quantum Theory of Gravitational Collapse (Lecture Notes on Quantum Conchology), Lect. Notes Phys. **631**, 255–299 (2003)
http://www.springerlink.com/

dependent degrees of freedom of the gravitational field (uniqueness theorems!), and these degrees of freedom have no proper quantum character of their own.

The model we work with is a spherically symmetric thin shell. Historically, the first physicist to use shells in quantum theory seems to be Dirac [1]. We shall try to do justice to the model and to take it absolutely seriously. There is no point in rushing towards prejudiced conceptions disregarding all kinds of problems that may emerge on the way.

These lecture notes are to explain in a coherent way a number of ideas scattered in various papers, but they also contain some new results. Section 2 gives a general account of some gauge-invariant methods in the canonical theory of generally covariant systems, whose spiritual fathers are Dirac, Bergmann, Kuchař, and Rovelli. New seems to be the observation that many conceptual and practical problems admit a solution if one restricts oneself to asymptotically flat models. This is a natural framework for the gravitational collapse. The asymptotic structure defines a decomposition of the full diffeomorphism group into the gauge group, the group of asymptotic symmetries and the rest that can be ignored. A physically meaningful choice of a complete set of gauge-invariant quantities – the so-called Dirac observables – is also provided by the structure. The asymptotic symmetries determine a gauge-invariant dynamics. In this way, some aspects of the notorious problem of time are dealt with.

Section 3 starts with an application of the methods to the model of a single, spherically symmetric, thin shell of light-like matter surrounded by its own gravitational field. The study of the space of the corresponding solutions to Einstein equations leads to the identification of a complete set of Dirac observables. Thus, the basic observables are found independently of the asymptotic structure in this (lucky) case, but they can still be considered as asymptotic properties of the system. The canonical theory of the shell due to Louko, Whiting, and Friedman is utilized to calculate the Poisson brackets of the observables and to find the generators of the asymptotic symmetries. The main new point is a careful description of asymptotic frames.

Finally, a quantum theory of the model is constructed in Sect. 4. The method of choice is the group-theoretical quantization because the phase space has a complicated boundary and the group quantization has been invented to deal with such a situation. The results are rather surprising: some quantum shells contract, cross their Schwarzschild radius inwards, bounce at the centre without creating a singularity, cross their Schwarzschild radius outwards, and re-expand to the infinity from which they originally came. This is no joke! It is possible because the quantum Schwarzschild radius is a mixture of black and white hole states. It is amusing that an old calculation [2] based on a very different technique has come to somewhat similar results. A couple of ideas are listed of how these results could be reconciled with the "observational evidence for black holes" in astrophysics. An interesting problem for future research arises: what is the nature of quantum geometry?

2 Gauge-Invariant Method in the Canonical Theory of Generally Covariant Systems

The present section attempts to give a general account of some new methods that have been developed in a couple of recent years. It is also a description of a project. It contains many claims without proofs. Those of these claims that make sense for finite dimensional systems have been proved and discussed extensively in [3] and [4] for general finite dimensional systems. The rest of the claims without proofs have hitherto been proved only for a few special cases, one of which will be described in the next section.

2.1 Space of Solutions, Gauge Group, and Asymptotic Symmetries

Mathematically, any *generally covariant system* is constituted by a set Ψ of geometrical objects (fields, and submanifolds such as particles or shells) on a background manifold $\mathcal{M}$. The set Ψ always includes a Lorentzian metric g on $\mathcal{M}$; it is sufficient to consider Ψ as consisting just of g in order to understand all manipulations with Ψ in this lecture. Details about complicated systems are explained in [5] and [6]. The background $\mathcal{M}$ has just a topological and differential structure (bare manifold without boundary) and has the meaning of a *spacetime* manifold. It can be of any dimension, but we assume here that it is four-dimensional. Several different backgrounds can be needed for one generally covariant system. For example, all three-dimensional manifolds that admit a Riemannian metric can play the role of Cauchy surfaces in general relativity; different topologies of Cauchy surface define different background manifolds.

The object Ψ has to satisfy a dynamical equation on $\mathcal{M}$ (such as Einstein equations for g). We allow for a domain $\mathcal{D}_\Psi$ of Ψ in $\mathcal{M}$ where the solution is well defined to be a proper subset of $\mathcal{M}$. General covariance is the following property. Let Diff$\mathcal{M}$ be the diffeomorphism group of $\mathcal{M}$ and let $(\mathcal{M}, \Psi)$ be a spacetime that solves the dynamical equation. Then $(\mathcal{M}, \varphi_*\Psi)$ is another solution for any $\varphi \in$ Diff$\mathcal{M}$. The symbol φ_* denotes the push forward of all geometrical objects collected within Ψ by φ. For example, let $\mathcal{M}$ admit a global coordinate chart and let $\{X^\mu\}$ denote some global coordinates on $\mathcal{M}$; the action of a diffeomorphism φ on a metric field $g_{\mu\nu}(X)$ has the following result:

$$(\varphi_* g)_{\mu\nu}(X) = g_{\rho\sigma}\big(Y(X)\big)\frac{\partial Y^\rho}{\partial X^\mu}\frac{\partial Y^\sigma}{\partial X^\nu} , \qquad (1)$$

where the functions $Y^\rho(X)$ are defined by φ and define the action of φ in terms of the coordinates $\{X^\mu\}$:

$$Y^\rho(X) = X^\rho\big(\varphi^{-1}(X)\big) . \qquad (2)$$

The map φ_* is an "active" transformation so that e.g. $\varphi_* g$ is a different field on $\mathcal{M}$ than g is, in general. If Ψ and φ are such that $\varphi_*\Psi = \Psi$, then we say that Ψ admits a *spacetime symmetry* φ. The reason why we prefer to work with active transformations rather than coordinate transformations is a technical one: active transformations can be globally defined even if there are no global coordinate charts. In consequence of that, they form groups so that the whole powerful apparatus of group theory can be applied.

We shall restrict ourselves to the special case of systems with *asymptotically flat* solutions. There are coordinate free methods (due to Penrose) to decide whether or not a spacetime is asymptotically flat (see, e.g., [7] or [8]). Then there is a universal *asymptotic boundary* $\partial_{\mathrm{as}}\mathcal{M}$ carrying an *asymptotic structure* that can be considered as common to all such solutions. For example, asymptotically flat solutions to Einstein equations all have the same scri, which has a geometrical structure that is richer than just the topological and differential one. This structure posses a symmetry group that we call G_{as}[1]. We shall not go into detail here, but we shall give a more complete account for an example later. The asymptotic structure enables to select *standard asymptotical frames of reference* at $\partial_{\mathrm{as}}\mathcal{M}$ that are also common to all solutions and are inhabitated by *asymptotic observers.*

One can try to give a more precise account of these ideas as follows. Let $\bar{\Gamma}$ be the space of all asymptotically flat solutions on $\mathcal{M}$. We assume that there is a subspace $\tilde{\Gamma} \subset \bar{\Gamma}$ and a subgroup G of Diff$\mathcal{M}$ with the following properties.

1. For each $\Psi \in \tilde{\Gamma}$, the spacetime $(\mathcal{M}, \Psi)$ can be conformally extended to $(\bar{\mathcal{M}}_\Psi, \Psi)$ by attaching a scri $\mathcal{I}_\Psi$ to $\mathcal{M}$ and a neighbourhood $\tilde{\mathcal{I}}_\Psi$ of $\mathcal{I}_\Psi$ in $\bar{\mathcal{M}}_\Psi$. Then there is a diffeomorphism $\Phi_\Psi : \tilde{\mathcal{I}}_\Psi \mapsto \bar{\mathcal{M}}_0$, such that Φ_Ψ maps $\mathcal{I}_\Psi$ onto $\mathcal{I}$, where $\bar{\mathcal{M}}_0$ is the conformal compactification of Minkowski spacetime and $\mathcal{I}$ is the scri of $\bar{\mathcal{M}}_0$.
2. Let $\varphi \in G$, $\Psi \in \tilde{\Gamma}$, $\Psi' = \varphi_* \Psi$ and $U = \Phi_\Psi \tilde{\mathcal{I}}_\Psi \cap \Phi_{\Psi'} \tilde{\mathcal{I}}_{\Psi'}$. Then φ can be differentiably extended to $\mathcal{I}_\Psi$ and $\varphi_\Psi : U \mapsto U$ defined by $\varphi_\Psi = \Phi_{\Psi'} \circ \varphi \circ \Phi_\Psi^{-1}$ induces a map $\varphi_\Psi|_{\mathcal{I}}$ on $\mathcal{I}$ of Minkowski spacetime that preserves the asymptotic structure and is an element of the asymptotic symmetry group G_{as}, $\varphi_\Psi|_{\mathcal{I}} \in G_{\mathrm{as}}$.
3. For given $\varphi \in G$, $\varphi_\Psi|_{\mathcal{I}}$ is independent of Ψ. In this way, there is a well-defined map $\pi_{\mathrm{as}} : G \mapsto G_{\mathrm{as}}$, and we assume that π_{as} is onto. It follows that this map is a homeomorphism if it exists. Let us define the *gauge group* G_0 of the system to be the kernel of π_{as}. It follows that G_0 is a normal subgroup of G.
4. For each $\varphi \in G_0$ and $\Psi \in \tilde{\Gamma}$, the extension of φ to $\mathcal{I}_\Psi$ is an odd-parity supertranslation[2] on $\mathcal{I}_\Psi$.

If this procedure works, then it is surely non unique. For example, $\mathcal{I}$ can be chosen to lie in a "different corner" of $\mathcal{M}$. We suppose that the non-uniqueness can either be limited by some suitable additional requirements or that it will not manifest itself in physical results.

[1] For example, a symmetry of the geometrical structure of flat spacetime is an isometry of the spacetime; the symmetry group of it is the Poincaré group.

[2] Justification for this assumption must be taken from the canonical theory of the next subsection; this shows how tentative and preliminary status the present subsection has.

The quotient $\Gamma := \tilde{\Gamma}/G_0$ is the *physical phase space.* (To be a full-fledged phase space, it has to be equipped with a symplectic structure.) Let us denote the projection of the quotient by $\pi_{\text{sol}} : \tilde{\Gamma} \mapsto \tilde{\Gamma}/G_0$.

Let $\varphi_1 \in G \setminus G_0$. Consider the action of φ_1 on a point Ψ_1 of an orbit γ of G_0 in $\tilde{\Gamma}$. Clearly, $\varphi_{1*}\Psi$ lies in another orbit γ' and $\gamma' \neq \gamma$. Let $\Psi_2 \in \gamma$ and $\Psi_2 \neq \Psi_1$. Does $\varphi_{1*}\Psi_2$ lie in the same orbit as $\varphi_{1*}\Psi_1$? It does because G_0 is a normal subgroup of G. Hence φ_1 sends orbits into orbits. Further, let φ_2 be from the same class of G/G_0 as φ_1 is. Then φ_2 also sends γ into γ'. Since G_{as} can be identified with the factor group G/G_0, this shows that the asymptotic symmetry group acts on the physical phase space.

Consider a surface σ in $\tilde{\Gamma}$ that cuts each orbit exactly once. Such a surface is called a *section* of the quotient $\tilde{\Gamma}/G_0$. Each section of $\tilde{\Gamma}/G_0$ breaks the gauge group and can, therefore, be called *covariant gauge fixing.* A section σ can also be described as follows. Let X^μ be coordinates on $\mathcal{M}$ and o^A coordinates on Γ. Then, for the special case that $\Psi = g$, σ determines the set of functions

$$g_{\mu\nu}(o, X), \tag{3}$$

that is, metric components with respect to the coordinates X^μ at the point g of intersection between σ and the orbit given by the coordinates o^i. The coordinates X^μ can be arbitrary and the same section can so be represented by different sets of functions (3). This is why the gauge fixing is called "covariant." Observe also that no global coordinate chart is necessary to describe σ.

Transformations between two different gauge fixings is constituted by a set of diffeomorphisms $\varphi(o)$, one for each orbit γ determined by coordinates o^i. Such gauge transformations are common in every gauge theory. For example, in electrodynamics, one often uses the so-called Coulomb gauge. This is defined by a differential equation for the components of the potential. One also uses the so-called axial gauge that is defined by an algebraic equation for the potential. The transformation between these two gauges must, therefore, depend on the potential in a rather complicated and non-local way. Quantum field theory constructed in a given gauge *cannot* be made invariant with respect to all field dependent gauge transformations. One hopes to overcome this difficulty by working with gauge-invariant quantities.

Similarly, for generally covariant systems, the two quantum theories constructed in two different gauges that are related by a field dependent transformation are not unitarily equivalent, unless they are limited to relations between Dirac observables [43].

Let $o \in \Gamma$ and $\Psi(o) = \sigma \cap \pi_{\text{sol}}^{-1} o$. It is a simple exercise to show that there is an element φ_o in each class of G/G_0 such that $\varphi_{o*}\Psi(o) \in \sigma$; if $\Psi(o)$ does not admit any spacetime symmetry, then φ_o is even unique! Thus, given a gauge fixing, we find a (o-dependent) extension of each asymptotic symmetry to the whole of $\mathcal{M}$. Even if gauge (and, in general, o-) dependent, this extension is a practical tool that will be used later.

2.2 Phase Space

We assume now that $\mathcal{M}$ has the structure $\Sigma\times\mathbb{R}$ and that the dynamical equation admits a well-posed Cauchy problem on the manifold Σ. Let us denote a Cauchy datum for the solution Ψ by $(\Psi_\Sigma, \Pi_\Sigma)$, where Ψ_Σ and Π_Σ are some geometric objects on Σ. We assume that a Riemannian metric ${}^3g_{kl}$ with the meaning of the first fundamental form of Σ belongs to Ψ_Σ and a symmetric tensor field K_{kl} with the meaning of the second fundamental form of Σ belongs to Π_Σ. We assume further that $(\Sigma,\ {}^3g)$ is asymptotically Euclidean. This involves the existence of a special coordinate patch in Σ such that the corresponding components of all objects within Ψ_Σ and Π_Σ satisfy suitable fall-off asymptotic conditions, so-called *Cauchy data asymptotic conditions* (CDAC). Different examples of CDAC are given in [9] and in [8], Sect. 5.4.

For generally covariant systems, the objects Ψ_Σ and Π_Σ must satisfy some particular conditions (mostly differential equations) in order to constitute an initial datum for a solutions. These equations are called *constrains.*

The *phase space* $\mathcal{P}$ of the system is a manifold, the points of which are all unconstrained Cauchy data. It carries a symplectic structure $\bar{\Omega}$, which defines Poisson brackets. We assume that such a space can be constructed. For example, ${}^3g_{kl}$ and

$$\pi^{kl} := \sqrt{\mathrm{Det}({}^3g)}({}^3g^{kl}\,{}^3g^{rs} - {}^3g^{kr}\,{}^3g^{ls})K_{rs} \tag{4}$$

are canonically conjugated variables for the case that $\Psi = g$ [10].

We assume further that all constrained data form a submanifold $\mathcal{C}$ of $\mathcal{P}$, called *constraint surface*: each point of $\mathcal{C}$ is a Cauchy datum for a solution and each Cauchy datum for a solution lies in $\mathcal{C}$.

There is an important relation "$\sim$" between points of $\mathcal{C}$ that can be defined as follows. The points p_1 and p_2 are said to satisfy $p_1 \sim p_2$ if the solutions Ψ_1 and Ψ_2 determined by them satisfy

$$\Psi_2 = \varphi_*\Psi_1 \tag{5}$$

for some $\varphi \in G_0$. We assume that "$\sim$" is an equivalence relation. (For general relativity, this has been shown in [11]). The equivalence class of points at $\mathcal{C}$ are called *c-orbits* and denoted by λ. Each c-orbit determines a class of G_0-equivalent solutions and vice versa. Hence, the quotient $\mathcal{C}/\lambda$ is the physical phase space defined in Sect. 2.1:

$$\Gamma = \mathcal{C}/\lambda. \tag{6}$$

The quotient projection of the constraint surface $\mathcal{C}$ to the physical phase space Γ will be denoted by π_{phs}.

Let the constraint functionals

$$\mathcal{H}[\Psi_\Sigma, \Pi_\Sigma; x),\ \mathcal{H}_k[\Psi_\Sigma, \Pi_\Sigma; x), \tag{7}$$

where $k = 1, 2, 3$ and $x \in \Sigma$, define the constraint surface $\mathcal{C}$ by $\mathcal{H} = 0$ and $\mathcal{H}_k = 0$. It seems that they can be chosen for all generally covariant systems so that they obey the so-called Dirac algebra [12], see also [13]: Let x^k be coordinates, $\mathcal{N}(x)$ a scalar and $\mathcal{N}^k(x)$ a vector fields on Σ; the fields $\mathcal{N}(x)$ and $\mathcal{N}^k(x)$ are called *lapse* and *shift*, respectively. Let $\boldsymbol{\mathcal{N}} := (\mathcal{N}, \mathcal{N}^1, \mathcal{N}^2, \mathcal{N}^3)$ and

$$\mathcal{H}[\boldsymbol{\mathcal{N}}] := \int_\Sigma d^3x \big(\mathcal{N}(x)\mathcal{H}(x) + \mathcal{N}^k(x)\mathcal{H}_k(x)\big) \ . \tag{8}$$

Then

$$\{\mathcal{H}[\boldsymbol{\mathcal{N}}_1], \mathcal{H}[\boldsymbol{\mathcal{N}}_2]\} = \mathcal{H}[\boldsymbol{\mathcal{N}}] \ , \tag{9}$$

where

$$\mathcal{N} \ = \mathcal{N}_1^k \partial_k \mathcal{N}_2 - \mathcal{N}_2^k \partial_k \mathcal{N}_1 \ , \tag{10}$$

$$\mathcal{N}^k = \mathcal{N}_1^l \partial_l \mathcal{N}_2^k - \mathcal{N}_2^l \partial_l \mathcal{N}_1^k + \ ^3g^{kl}(\mathcal{N}_1 \partial_l \mathcal{N}_2 - \mathcal{N}_2 \partial_l \mathcal{N}_1) \ . \tag{11}$$

This implies that the Poisson brackets of the constraints vanish at the constraint surface; such a constraint surface is called *first class* [14], [15]. However, (9) holds only if the lapse and shift fields satisfy some fall-off conditions (see [9]), called *gauge group asymptotic conditions* (GGAC). An unexpected result of [9] is that $\boldsymbol{\mathcal{N}}$ need not approach zero asymptotically in order to satisfy (GGAC) but can contain arbitrary odd-parity supertranslations.

The lapse and shift fields can even approach linear functions of coordinates asymptotically. This corresponds to infinitesimal Poincaré transformations at infinity (cf. [16] and [9]). The condition that $\boldsymbol{N} \to \boldsymbol{N}_\infty$, where $\boldsymbol{N}_\infty$ represents the asymptotic behaviour of $\boldsymbol{N}$ corresponding to an element of the Lie algebra of Poincaré group is called *Poincaré group asymptotic condition* (PGAC); the form of $\boldsymbol{N}_\infty$ as a function on Poincaré algebra is given in [16] and [9]. The corresponding functionals $\mathcal{H}[\boldsymbol{N}]$ are not differentiable (their variations lead to surface integrals at infinity) or are not even convergent. This can be improved by adding *surface terms* at infinity to them that we denote by $\mathcal{H}_\infty[\boldsymbol{N}_\infty]$: the functional form of $\mathcal{H}_\infty[\boldsymbol{N}_\infty]$ is given in [16] and [9]. An analysis of similar surface terms for the null-infinity is given in [8]. The expressions $\mathcal{H}[\boldsymbol{N}] + \mathcal{H}_\infty[\boldsymbol{N}_\infty]$ have finite values on constraint surface and generate asymptotic Poincaré transformation[3].

Observe that the CDAC form a part of the definition of the phase space, while GGAC and PGAC help define some transformations on this phase space. In particular, the CDAC must be preserved in all transformations satisfying the GGAC or PGAC, see [9]. In this way, the set of all lapse and shift fields is divided into those that are associated with infinitesimal gauge transformations, infinitesimal dynamical symmetries and the rest that is not interesting. Again, the gauge $\boldsymbol{N}$'s form a large subset of the symmetry $\boldsymbol{N}$'s. What is the relation between this Cauchy surface picture and the spacetime picture of Sect. 2.1? We believe that such a relation can be found using the technique invented by

[3] The surface terms may be uniquely determined (except for additive constants) by boundary conditions imposed on the dynamics of the fields at the surface. For a finite surface, this has been shown by [17].

Kuchař, the so-called Kuchař variables (embeddings, see Sect. 2.4). In particular, the CDAC must be related to asymptotic conditions imposed on the embeddings. An example of how this may work will be given in the next section.

If $\mathcal{N}$ satisfies the GGAC, the transformation generated by $\mathcal{H}[\mathcal{N}]$ in the phase space shifts the points of $\mathcal{C}$ along c-orbits, and any point of a given c-orbit can be reached in this way from any other point of it. This is the reason to call λ a c-orbit (constraints orbit).

For any first class constraint surface, there is a unique way of how the symplectic structure $\bar{\Omega}$ of the phase space $\mathcal{P}$ defines a symplectic structure Ω on the physical phase space Γ: First step is to pull back the two-form $\bar{\Omega}$ to $\mathcal{C}$. The result is a two-form $\tilde{\Omega}$. The form $\tilde{\Omega}$ is degenerated along c-orbits; it is not a symplectic form, but it is closed ($d\tilde{\Omega} = 0$, where d denotes the external differentiation). The second step is a proof that there is a unique symplectic form Ω on Γ such that the pull back of it by the projection π_{phs} to $\mathcal{C}$ is $\tilde{\Omega}$. In this way, we finally obtain the full-fledged physical phase space (Γ, Ω).

2.3 Observables and Dynamical Symmetries

The action of the gauge group has been quotiented away from the physical phase space, so it is a gauge-invariant structure. Functions on it are called *observables*:

$$o : \Gamma \mapsto \mathbb{R} \ . \tag{12}$$

The central idea of the gauge-invariant method described by these lectures is to answer all physically interesting questions by manipulating quantities defined on the physical phase space, such as observables. To interpret or justify such manipulations, we need the connection of Γ to the larger spaces $\mathcal{C}$, $\mathcal{P}$ and $\tilde{\Gamma}$.

Each observable o determines a function $\tilde{o}$ on $\mathcal{C}$ by

$$\tilde{o} := o \circ \pi_{\mathrm{phs}} \ . \tag{13}$$

Clearly, all such functions must be constant along c-orbits:

$$\{\tilde{o}, \mathcal{H}[\mathcal{N}]\} = 0 \tag{14}$$

for all $\mathcal{N}$ that satisfy GGAC; they are called *Dirac observables* [14]. Examples of Dirac observable are the quantities $\mathcal{H}[\boldsymbol{N}] + \mathcal{H}_\infty[\boldsymbol{N}_\infty]$ for all fields $\boldsymbol{N}$ that satisfy PGAC, see [9].

Let $\{o^i\}$ be some coordinates on Γ; they form a *complete set of observables*. Similarly, the corresponding functions $\tilde{o}^i$ form *complete set of Dirac observables*. One needs complete sets with some additional properties to quantize by a gauge-invariant way, and one needs still more observables to describe all interesting properties of the quantum system. This will be shown later by an example. The idea of basing quantization of gravity on complete sets of Dirac observables encounters a difficulty: there is no single quantity of this kind known in general relativity, for example in the case when the Cauchy surface is compact [18], [19]. Even if it were known, it would be non-local [20]. In the asymptotically flat case, however, two complete sets are known: they are associated with the asymptotic in- and out-fields. They have been described, within a perturbation theory, by DeWitt ("asymptotic invariants") [21], and for the exact theory by Ashtekar ("radiative modes") [22]. We propose to work out a theory for this favorable case first and then to see if anything can be done for the other cases.

Let $\bar{h} : \mathcal{P} \mapsto \mathcal{P}$ be a map that preserves the symplectic form $\bar{\Omega}$ (symplectomorphism) and the constraint surface $\mathcal{C}$ in $\mathcal{P}$. Such a map is called *extended dynamical symmetry* (extended to the extended phase space $\mathcal{P}$). We use the word "dynamical" to distinguish it from spacetime symmetries.

Clearly, such a map $\bar{h}$ induces a map of the constraint surface onto itself: $\tilde{h} := \bar{h}|_{\mathcal{C}}$ and $\tilde{h} : \mathcal{C} \mapsto \mathcal{C}$. The most important properties of $\tilde{h}$ are: it preserves the form $\tilde{\Omega}$ and maps the c-orbits onto c-orbits. Such a map $\tilde{h}$ is called *dynamical symmetry*. Each dynamical symmetry $\tilde{h}$ induces a map $h : \Gamma \mapsto \Gamma$ that preserves the symplectic structure Ω, so h is a symplectomorphism of the physical phase space (Γ, Ω).

An infinitesimal extended dynamical symmetry $\overline{dh}$ is generated via Poisson brackets by a function $-\bar{H}dt : \mathcal{P} \mapsto \mathcal{P}$ (the sign is chosen for later convenience). The restriction $\tilde{H}$ of $\bar{H}$ to $\mathcal{C}$ is a Dirac observable and so it defines a unique function H on the physical phase space Γ. The function $-Hdt$ generates the infinitesimal symplectomorphism dh in Γ that is induced by $\overline{dh}$, for proofs, see [9], [3] and [4]. (We denote infinitesimal maps by a symbol composed from "d" and a letter; the letter itself has no further meaning.)

The infinitesimal transformations generated by $(\mathcal{H}[\boldsymbol{N}] + \mathcal{H}_\infty[\boldsymbol{N}_\infty])dt$ are dynamical symmetries for all $\boldsymbol{N}$ satisfying PGAC. Their Poisson brackets at the constraint surface form the Lie algebra of Poincaré group [9]. They have in general non-vanishing Poisson brackets with other Dirac observables generating a non-trivial dynamics for them. We shall give an interpretation of this dynamics in the next two subsections.

2.4 Transversal Surfaces

A transversal surface $\mathcal{T}$ is a submanifold of $\mathcal{C}$ that intersects each c-orbit exactly once and transversally; $\mathcal{T}$ is a section of the quotient $\mathcal{C}/\lambda$. Each transversal surface inherits a two-form $\Omega_{\mathcal{T}}$ from $\mathcal{C}$ by pull back of $\tilde{\Omega}$ to $\mathcal{T}$ by the injection map of $\mathcal{T}$ into $\mathcal{C}$. It is not only closed but also non-degenerate, so $(\mathcal{T}, \tilde{\Omega})$ is a symplectic manifold (see also [9], where transversal surface is called "gauge condition").

The projection π_{phs} maps $\mathcal{T}$ to Γ and its restriction $\pi_{\text{phs}}|_{\mathcal{T}}$ is actually a bijection having an inverse. The map $\pi_{\text{phs}}|_{\mathcal{T}}$ can be shown to be a symplectomorphism of the spaces $(\mathcal{T}, \Omega_{\mathcal{T}})$ and (Γ, Ω). Thus, each transversal surface is a "model" of the physical phase space. If we wish to calculate Ω in the coordinates $\{o^i\}$, then we just have to calculate $\Omega_{\mathcal{T}}$ in the coordinates $\{\tilde{o}^i\}$.

An important issue is the relation between a transversal surface $\mathcal{T}$ and Cauchy surfaces in solution space-times. Such a relation is provided by a gauge fixing σ. As it has been shown above, σ can be represented as a set of functions $\Psi(o, X)$, where $\{o^i\}$ are coordinates on Γ and X^μ coordinates on $\mathcal{M}$. Consider the spacetime $(\mathcal{M}, \Psi(o,X))$; a Cauchy surface in such a spacetime can be described by an embedding $\hat{X}(o) : \Sigma \mapsto \mathcal{M}$ given by the set of functions $X^\mu(o,x)$. Having these functions and the solution $\Psi(o,X)$, we can calculate the corresponding Cauchy datum $(\Psi_{\hat{X}(o)}, \Pi_{\hat{X}(o)})$ and so determine a point at $\mathcal{C}$. More specifically, this point lies at the c-orbit $\lambda(o)$ corresponding to the point $o \in \Gamma$: $\lambda(o) = \pi_{\text{phs}}^{-1} o$. In this way, there is a unique point $(\Psi_{\hat{X}(o)}, \Pi_{\hat{X}(o)})$ at $\mathcal{C}$ for each λ, and this defines a transversal surface $\mathcal{T}$.

Let us denote the resulting map that sends the set of embeddings $\{\hat{X}(o)\}$ to the transversal surface $\mathcal{T}$ by χ_σ:

$$\chi_\sigma\{\hat{X}(o)\} = \mathcal{T} . \tag{15}$$

The functions $\Psi(o,X)$ and $X(o,x)$ must be sufficiently smooth or else $\mathcal{T}$ would be rather jumpy.

One can show that χ_σ can even be a bijection between the space of embedding sets $\{X(o,x)\}$ and the space of transversal surfaces $\mathcal{T}$. The condition is that the solutions $\Psi(o,X)$ do not admit any spacetime symmetry [5]. Then each transversal surface and σ determine together a set of embeddings $\{X(o,x)\}$, one embedding for each solution spacetime $(\mathcal{M}, \Psi(o,X))$:

$$\{X(o,x)\} = \chi_\sigma^{-1}\mathcal{T} . \tag{16}$$

One can say that, given a gauge, each transversal surface defines a *many-finger-time* instant in each solution spacetime. Transversal surfaces can, therefore, also be called many-finger-time levels.

An interesting point is that the variables o^i and $\{X(o,x)\}$ – coordinates on the physical phase space and a set of embeddings, one for each solution – form a coordinate system on $\mathcal{C}$ if a gauge σ is specified. They are called generalized[4] *Kuchař variables*, and they constitute a useful tool for many calculations because they provide a neat division of variables into gauge, physical and dependent degrees of freedom on one hand and a bridge between the four-dimensional and the three-dimensional pictures on the other.

2.5 Time Evolution

Dirac observables are constant along whole solutions and so they are not only gauge invariant but a kind of integrals of motion. It seems that their dynamics is trivial: they just stay constant. Bergmann [18] characterized that as “frozen dynamics”.

However, as early as 1949, Dirac [25] has put forward a theory of time evolution that is based on transversal surfaces and symmetries and that leads to a non trivial time evolution of Dirac observables. The idea is that a physically sensible evolution results if the “pure” dynamics (such as staying constant) is compared with a fiducial “zero” dynamics defined by a symmetry. The account presented here is a generalization [3], [4] of Dirac theory. It can also be understood as an

[4] The embeddings introduced originally by Kuchař [23], [24] were o-independent.

illustration of the discovery [26] that the Hamiltonian dynamics needs a frame of reference and that different frames lead to differently looking time evolutions of one and the same system, see also [8].

Let $\tilde{h}$ be dynamical symmetry and $\mathcal{T}$ a transversal surface. Then $\tilde{h}\mathcal{T} = \mathcal{T}'$ is another transversal surface and $\tilde{h}$ is a symplectomorphism between the spaces $(\mathcal{T}, \Omega_{\mathcal{T}})$ and $(\mathcal{T}', \Omega_{\mathcal{T}'})$. Thus, $\tilde{h}$ can shift many-finger-time levels.

Suppose a one-dimensional group of extended dynamical symmetries $\bar{h}(t)$ is given; t is the parameter of the group, $\bar{h}(t_1)\cdot\bar{h}(t_2) = \bar{h}(t_1+t_2)$. Let the generator of the group be $-\bar{H}$, and let the induced groups and generators on $\mathcal{C}$ and Γ be denoted by $\tilde{h}(t)$, $h(t)$ and $-H$. We use the group $\tilde{h}(t)$ to build up a reference system in $\mathcal{C}$ with respect to which we shall describe the motion represented by c-orbits in a similar way as the particle world lines in Minkowski spacetime are described by their coordinates in an inertial frame.

Let $\mathcal{T}$ be a transversal surface. Define $\mathcal{T}_t := \tilde{h}(t)\mathcal{T}_0$. These surfaces form a one-dimensional family that we denote by $\{\mathcal{T}_t\}$. The family determines a subset $\eta_\lambda(t)$ of any c-orbit λ by

$$\eta_\lambda(t) := \mathcal{T}_t \cap \lambda \ . \tag{17}$$

The curve $\eta_\lambda(t)$ lies in λ and is one-dimensional in spite of λ being itself many (or even infinitely many) dimensional. We call $\eta_\lambda(t)$ a trajectory of the system with respect to the family $\{\mathcal{T}_t\}$.

Let $\tilde{o}_0$ be a Dirac observable representing some measurement done at the time level[5] $\mathcal{T}_0$. The values of the function $\tilde{o}_0|_{\mathcal{T}_0}$ on $\mathcal{T}_0$ give results of measurement at each point of $\mathcal{T}_0$, that is, at each instantaneous state of the system.

We define *the same measurement* at the time $\mathcal{T}_t$ to be represented by Dirac observable $\tilde{o}_t := \tilde{o}_0 \circ \tilde{h}(-t)$. The function $\tilde{o}_t$ is the image of $\tilde{o}_0$ by the map $\tilde{h}(t)$: it gives the same results at the states that are related by the symmetry:

$$\tilde{o}_t\big(\tilde{h}(t)p\big) = \tilde{o}_0(p) \tag{18}$$

for any point $p \in \mathcal{C}$. Actually, the set $\{\tilde{o}_t | t \in \mathbb{R}\}$ of Dirac observables is a special case of Rovelli's "evolving constants of motion" [27].

Now, we are ready to define the time evolution.

Definition

Time evolution is the change in the results of the same measurement done at different times along a dynamical trajectory of the system.

Clearly, these results are given by the function $\tilde{o}_t\big(\eta_\lambda(t)\big)$ for the measurement represented by $\{\tilde{o}_t | t \in \mathbb{R}\}$ and the trajectory λ. The function can be written in two ways.

[5] Indeed, a given Dirac observable cannot, in general, be measured anywhere on the constraint surface but only in a neighbourhood of a transversal surface that belongs to the definition of the observable, see [4].

Schrödinger Picture. By substituting from the definition of $\tilde{o}_t$, we have

$$\tilde{o}_t\big(\eta_\lambda(t)\big) = \tilde{o}_0\big(\tilde{h}(-t)\eta_\lambda(t)\big) \ . \tag{19}$$

Let us define

$$\tilde{\xi}_\lambda(t) := \tilde{h}(-t)\eta_\lambda(t) \ ; \tag{20}$$

$\tilde{\xi}_\lambda(t)$ is a curve in $\mathcal{T}_0$ and can be viewed as a "time-dependent state of the system." Then we can write

$$\tilde{o}_t\big(\eta_\lambda(t)\big) = \tilde{o}_0|_{\mathcal{T}_0}\big(\tilde{\xi}_\lambda(t)\big) \ . \tag{21}$$

The measurement is represented by a single (time-independent) Schrödinger observable $\tilde{o}_0|_{\mathcal{T}_0}$.

Heisenberg Picture. Since $\tilde{o}_t$ is a Dirac observable for each t, it is constant along c-orbits, and we find

$$\tilde{o}_t\big(\eta_\lambda(t)\big) = \tilde{o}_t\big(\eta_\lambda(0)\big) \ . \tag{22}$$

All quantities can again be taken at $\mathcal{T}_0$:

$$\tilde{o}_t\big(\eta_\lambda(t)\big) = \tilde{o}_t|_{\mathcal{T}_0}\big(\eta_\lambda(0)\big) \ . \tag{23}$$

Now, the state is described by a single point $\eta_\lambda(0)$ at $\mathcal{T}_0$ (time independence) while the measurement is described by a set of functions $\tilde{o}_t|_{\mathcal{T}_0}$ on $\mathcal{T}_0$, constituting a time-dependent Heisenberg observable $\{\tilde{o}_t|_{\mathcal{T}_0}|t \in \mathbb{R}\}$.

The dynamical equation for the Schrödinger state $\tilde{\xi}(t)$ or the Heisenberg observable $\{\tilde{o}_t|_{\mathcal{T}_0}|t \in \mathbb{R}\}$ can be calculated with the following results: The Schrödinger state $\tilde{\xi}(t)$ is an integral curve in $(\mathcal{T}_0, \Omega_{\mathcal{T}_0})$ of the Hamiltonian vector field $d\tilde{H}|^{\#}_{\mathcal{T}_0}$ of $\tilde{H}|_{\mathcal{T}_0}$:

$$\frac{d\tilde{\xi}_\lambda(t)}{dt} = d\tilde{H}|^{\#}_{\mathcal{T}_0} \ . \tag{24}$$

The Heisenberg observable $\{\tilde{o}_t|_{\mathcal{T}_0}\}$ must satisfy the differential equation

$$\frac{d\tilde{o}_t|_{\mathcal{T}_0}}{dt} = \{d\tilde{o}_t|_{\mathcal{T}_0}, \tilde{H}|_{\mathcal{T}_0}\}_{\Omega|_{\mathcal{T}_0}} \ , \tag{25}$$

where $\{\cdot,\cdot\}_{\Omega|_{\mathcal{T}_0}}$ is the Poisson bracket of $(\mathcal{T}_0, \Omega_{\mathcal{T}_0})$.

All definitions and equations hitherto written down depend on the choice of $\mathcal{T}_0$. This leads to the impression that our dynamics depends not only on the chosen symmetry group $\tilde{h}(t)$ but also on $\mathcal{T}_0$. This impression may be strengthened if one reads Dirac's paper. However, Dirac's choice of the group was related to his choice of the surface. Here, these two choices are independent.

Actually, the exposition given as yet has aimed at some physical interpretation of the final result, which comes only now: The symplectomorphism $\pi_{\text{phs}}|_{\mathcal{T}_0}$ maps the equations (24) and (25) at $\mathcal{T}_0$ to the following equations at Γ,

$$\frac{d\xi_\lambda}{dt} = dH^\# \tag{26}$$

and

$$\frac{do_t}{dt} = \{o_t, H\}_\Omega \ , \tag{27}$$

where

$$\xi_\lambda(t) := \pi_{\text{phs}}|_{\mathcal{T}_0}\tilde{\xi}_\lambda(t) \ , \quad \tilde{o}_t|_{\mathcal{T}_0} = o_t \circ \pi_{\text{phs}}|_{\mathcal{T}_0} \ . \tag{28}$$

Equations (26) and (27) have the same form for any choice of $\mathcal{T}_0$ and depend only on the group $h(t)$.

The symmetry group $h(t)$ is to be chosen as a suitable subgroup of the asymptotic symmetries. To be sure, asymptotic symmetries cannot be extended to the whole of $\mathcal{M}$ in a unique way (a gauge-dependent extension has been constructed at the end of Sect. 2.1). However, the action of such extended map on a gauge-invariant quantity does not depend on the way it has been extended. The dynamics constructed in this way is, therefore, gauge invariant and it seems to be physically sound.

3 A Model: Gravitating Shell

This section will illustrate the notions and methods introduced in the preceding one. At the same time, the dynamics of the shell will be given a form that will be suitable for quantization.

3.1 Space of Solutions, Gauge Group, and Asymptotic Symmetries

The model consists of a spherically symmetric thin shell of light-like matter surrounded by its own gravitational field. Hence, there are two geometrical objects in $\bar{\Psi}$: a metric $g_{\mu\nu}$ describing the gravitational field and a three-dimensional light-like surface $\bar{S}$ that is a trajectory of the shell. The background manifold $\bar{\mathcal{M}}$ is $\mathbb{R}^3 \times \mathbb{R}$, where $\mathbb{R}^3$ is the manifold of Cauchy surface. The dynamical equations are Einstein's and matter equations. Any solution $\bar{\Psi}$ can be constructed by sticking together a piece of Schwarzschild solution of mass parameter M with the meaning of gravitational radius, and a piece of Minkowski spacetime, along a spherically symmetric null hypersurface so that the points with the same radius coordinate cover each other (see, e.g., [6]).

The spherical symmetry enables us to reduce the dimension of the problem by two. The effective spacetime can be considered as the space of the rotation group orbits; the background manifold $\mathcal{M}$ is $\mathbb{R} \times \mathbb{R}_+$, where $\mathbb{R}_+ := (0, \infty)$. The four-dimensional metric decomposes into a two-dimensional metric g_{AB} on $\mathcal{M}$ and a scalar field R on $\mathcal{M}$ (radius of the rotation group orbit). The shell trajectory becomes a curve S. These are the three geometrical objects in Ψ.

Luckily, for this simple model, all solutions can be explicitly written down. Thus, we can make a list of all physically different solutions to a basis of our analysis. The list can best be given in a particular gauge determined by a covariant gauge fixing σ. To specify σ, we choose the coordinates U and V on $\mathcal{M}$ with ranges $U \in \mathbb{R}$ and $V \in (U, \infty)$. The boundary of $\mathcal{M}$ (that does not belong to it) consists of three pieces: $\partial_0\mathcal{M}$ given by the equation $V = U$, with a coordinate $T_0 := (U+V)/2$, $T_0 \in \mathbb{R}$, $\partial_+\mathcal{M}$ defined by the limit $V \to \infty$ and described by coordinate $U_\infty = U$, $U_\infty \in \mathbb{R}$, and $\partial_-\mathcal{M}$ defined by the limit $U \to -\infty$ and described by coordinate $V_\infty = V$, $V_\infty \in \mathbb{R}$. With respect to the fixed coordinates U and V, we set conditions on the components of the representative metric, scalar field and shell trajectory as follows (for more detail and motivation, see [6] and [28]).

1. The coordinates U and V are double null coordinates:
$$ds^2 = -A(U,V)dUdV \ . \tag{29}$$
2. For out-going shells, S is defined by $U = \text{const}$, and the coordinate U coincides with a retarded time at $\partial_+\mathcal{M}$.
3. For in-going shells, S is defined by $V = \text{const}$, and the coordinate V coincides with an advanced time at $\partial_-\mathcal{M}$.
4. The functions $A(U,V)$ and $R(U,V)$ are continuous at the shell.

Then the representative solutions can be described as follows (see [6] and [28]). We introduce a parameter η with two values $+1$ and -1 to define the direction of radial motion of the shell: expanding for $\eta = +1$ and contracting for $\eta = -1$.

Out-Going Shell. ($\eta = +1$).

1. The shell trajectory is given by
$$U = w \ , \tag{30}$$
$w \in \mathbb{R}$ being a measurable parameter: the retarded arrival time of the out-going shell.
2. Left from the shell, $U > w$, the two-metric $A(U,V)$ and the scalar field $R(U,V)$ are given by
$$A = 1 \ , \quad R = \frac{-U+V}{2} \ . \tag{31}$$
3. Right from the shell, $U < w$, they are given by
$$A = \frac{1}{\kappa(f_+)e^{\kappa(f_+)}}\frac{-w+V}{4M}\exp\left(\frac{-U+V}{4M}\right) \ , \tag{32}$$
and
$$R = 2M\kappa(f_+) \ , \tag{33}$$

where

$$f_+ = \left(\frac{-w+V}{4M} - 1\right) \exp\left(\frac{-U+V}{4M}\right) . \tag{34}$$

The Kruskal function κ is defined by its inverse

$$\kappa^{-1}(x) = (x-1)e^x \tag{35}$$

on the interval $x \in (0, \infty)$.

At the shell,

$$\lim_{U \to w_-} f_+ = \kappa^{-1}\left(\frac{-w+V}{4M}\right) , \tag{36}$$

hence, we have $R \to (-w+V)/2$ and $A \to 1$ and the metric is continuous, but its derivatives have a jump. In this way, these coordinates carry the differential structure C^1 that is determined by the metric. The stress-energy tensor of the shell matter can be calculated by the formula given in [29]. Let us denote this solution by $\Psi_+(M, w; U, V)$.

The boundary $R \to 0$ is given by $V \to U$ for $U > w$ and by $f_+ \to -1$ or

$$V \to w + 4M\kappa \left[-\exp\left(\frac{U-w}{4M}\right)\right] \tag{37}$$

for $U < w$. The curve defined by the right-hand side of (37) is space-like running from the point $(U, V) = (w, w)$ to the point $(U, V) = (-\infty, w + 4M)$. It is the boundary of the open domain $\mathcal{D}_\Psi$ in $\mathcal{M}$ where the representative solution $(\mathcal{M}, \Psi_+)$ is well defined. The curve $V = w + 4M$ is the horizon, $R = 2M$.

In-Going Shell. $(\eta = -1)$.

1. The shell trajectory is given by

$$V = w \; ; \tag{38}$$

   $w \in \mathbb{R}$ is the advanced departure time of the in-going shell.
2. Left from the shell, $V < w$, the two-metric $A(U, V)$and the scalar field $R(U, V)$ are given by

$$A = 1 \; , \quad R = \frac{-U+V}{2} . \tag{39}$$

3. Right from the shell, $V > w$, they are given by

$$A = \frac{1}{\kappa(f_-)e^{\kappa(f_-)}} \frac{-U+w}{4M} \exp\left(\frac{-U+V}{4M}\right) , \tag{40}$$

   and

$$R = 2M\kappa(f_-) , \tag{41}$$

   where

$$f_- = \left(\frac{-U+w}{4M} - 1\right) \exp\left(\frac{-U+V}{4M}\right) . \tag{42}$$

Let us denote this solution by $\Psi_-(M, w; U, V)$. Again, the metric is continuous at the shell. The boundary of the domain $\mathcal{D}_\Psi$ is given by the equation $f_- = -1$. Observe that the solutions $\Psi_-(M, w; U, V)$ fill in the lower corner of $\mathcal{M}$ and so every point of the manifold $\mathcal{M}$ is used by both solutions $\Psi_+(M, w; U, V)$ and $\Psi_-(M, w; U, V)$.

The first employment of our list is the specification of the physical phase space. It has two components, Γ_+ and Γ_- corresponding to $\eta = +1$ and $\eta = -1$. They have both the manifold structure of $\mathbb{R}_+ \times \mathbb{R}$, where $M \in \mathbb{R}_+$ and $w \in \mathbb{R}$. M and w can serve as coordinates on Γ_η. It is not yet clear that these coordinates are regular; this ought to follow from the regularity of the symplectic form of the physical phase space written with respect of these coordinates. The symplectic form cannot be inferred from our list of solutions and will be calculated later.

The group of general four-dimensional diffeomorphisms can also be reduced by the symmetry: we admit only diffeomorphisms that commute with rotations so that the rotation group orbits are mapped onto such orbits. Clearly, regular centre points are sent only to regular centre points – so that the regular centre is invariant – while a spherical orbit can be sent to any spherical orbit. Hence, the reduced group is Diff$\mathcal{M}$ for our two-dimensional background manifold $\mathcal{M}$.

The second employment of the solution list given above is to find symmetries of the system. The symmetries will help us to decompose the group Diff$\mathcal{M}$, to define the most important observables and to construct the dynamics. To find the symmetries, we observe that some pairs of our representative solutions are isometric or conformally related to each other. For each such pair, there is a unique element of Diff$\mathcal{M}$ that implements the relation.

Time Shifts. Clearly, the solutions $\Psi_\eta(M, w; U, V)$ and $\Psi_\eta(M, w + t; U, V)$ are isometric for any $t \in \mathbb{R}$, M, w and η. The corresponding diffeomorphism $\varphi_H(t)$ is given by

$$\varphi_H(t) : (U, V) \mapsto (U + t, V + t) \tag{43}$$

For example, if $\eta = +1$, the shell is sent to $U = w + t$, the new metric and the scalar are, for $U > w + t$,

$$A = 1 \ , \quad R = \frac{-U + V}{2} \ , \tag{44}$$

while, for $U < w + t$,

$$A = \frac{1}{\kappa(f'_+) e^{\kappa(f'_+)}} \frac{-w - t + V}{4M} \exp\left(\frac{-U + V}{4M}\right) \ , \tag{45}$$

and

$$R = 2M\kappa(f'_+) \ , \tag{46}$$

where

$$f'_+ = \left(\frac{-w - t + V}{4M} - 1\right) \exp\left(\frac{-U + V}{4M}\right) \ . \tag{47}$$

Hence, these are just the fields (31)–(34) with $w \mapsto w+t$. The proof for $\eta = -1$ is analogous. Observe that the map $\varphi_H(t)$ is well defined everywhere on $\mathcal{M}$ and independent of the parameters η, M and w. Hence, all such maps form a one-dimensional group G_H with parameter t. The group acts on the physical phase space: $\varphi_H(t) : (M, w) \mapsto (M, w+t)$. Finally, its action can be uniquely extended to the boundary:

$$U_\infty \mapsto U_\infty + t \ , \quad V_\infty \mapsto V_\infty + t \ , \quad T_0 \mapsto T_0 + t \ . \tag{48}$$

Dilatations. Clearly, the solutions $\Psi_\eta(M, w; U, V)$ and $\Psi_\eta(e^s M, e^s w; U, V)$ are conformally related for all $s \in \mathbb{R}$, η, M and w. On the other hand, they are sent into each other by a composite map $\Big(\varphi_D(s), \omega\big(\varphi_D(s)\big)\Big)$, where

$$\varphi_D(s) : (U, V) \mapsto (e^s U, e^s V) \tag{49}$$

is a diffeomorphism and

$$\omega\big(\varphi_D(s)\big) : g_{\mu\nu} \mapsto e^{2s} g_{\mu\nu} \tag{50}$$

is a constant conformal deformation (CCD) that acts only on the (four-dimensional) metric: it scales the two-dimensional metric g_{AB} by e^{2s} and the scalar field R by e^s.

Actually, the dilatation is nothing but a rescaling: all quantities with a dimension are rescaled by the corresponding power of e^s. The coordinates U and V have acquired their scale from the metric through the gauge fixing. The Newton constant G must also be rescaled: $\mathrm{G} \mapsto e^{2s}\mathrm{G}$!

To see that the map (49) makes the job, let us restrict ourselves to the case $\eta = +1$ and study the action of the map on the objects in the spacetime defined by (31)–(34). The metric is transformed by $\varphi_D(s)_*$ to

$$ds^2 = -e^{-2s} A(e^{-s}U, e^{-s}V) dU dV \ , \tag{51}$$

the scalar field to

$$R = R(e^{-s}U, e^{-s}V) \tag{52}$$

and the shell position changes as follows

$$w \mapsto e^s w \ . \tag{53}$$

Thus, for $U > e^s w$, the new metric A is e^{-2s} and the new scalar R is

$$e^{-s} \frac{-U+V}{2} \ ; \tag{54}$$

for $U < e^s w$, the new metric A is

$$A = \frac{e^{-2s}}{\kappa(f'_+) e^{\kappa(f'_+)}} \frac{-we^s + V}{4Me^s} \exp\left(\frac{-U+V}{4Me^s}\right) \tag{55}$$

and the scalar R,

$$R = 2M\kappa(f'_+) \ , \tag{56}$$

where

$$f'_+ = \left(\frac{-we^s + V}{4Me^s} - 1\right)\exp\left(\frac{-U+V}{4Me^s}\right) \ . \tag{57}$$

The subsequent CCD by e^{2s} brings A and R to

$$A = 1 \ , \quad R = \frac{-U+V}{2} \tag{58}$$

for $U > we^s$, to

$$A = \frac{1}{\kappa(f'_+)e^{\kappa(f'_+)}}\frac{-we^s + V}{4Me^s}\exp\left(\frac{-U+V}{4Me^s}\right) \tag{59}$$

and

$$R = 2Me^s\kappa(f'_+) \ , \tag{60}$$

for $U < we^s$, and the shell stays at $U = we^s$. But the fields (58)–(60) are just the same as (31)–(34) except for M being changed to Me^s and w to we^s.

We observe that the map $\Big(\varphi_D(s), \omega\big(\varphi_D(s)\big)\Big)$ is well defined on the whole $\mathcal{M}$ and independent of η, M and w. Since each CCD commutes with any diffeomorphism, the dilatations form a one-dimensional group G_D with parameter s. The part $\varphi_D(s)$ form themselves a group $G_{D\varphi}$ that is a subgroup of Diff$\mathcal{M}$. The dilatations act on the physical phase space:

$$\Big(\varphi_D(s), \omega\big(\varphi_D(s)\big)\Big) : (\eta, M, w) \mapsto (\eta, e^sM, e^sw) \ . \tag{61}$$

Finally, dilatations act on the boundary

$$U_\infty \mapsto e^sU_\infty \ , \quad V_\infty \mapsto e^sV_\infty \ , \quad T_0 \mapsto e^sT_0 \ . \tag{62}$$

The time shifts and dilatations generate a group. Each element of the group can be obtained in a unique way as a dilatation followed by a time shift. Let us write it in the form $\Big(\varphi_H(t), \varphi_D(s), \omega\big(\varphi_D(s)\big)\Big)$. If the action of the constituent maps on, say, the parameter w is taken into account, the following composition law can be found

$$\Big(\varphi_H(t_1), \varphi_D(s_1), \omega\big(\varphi_D(s_1)\big)\Big) \times \Big(\varphi_H(t_2), \varphi_D(s_2), \omega\big(\varphi_D(s_2)\big)\Big) = \\ \Big(\varphi_H(t_1 + e^{s_1}t_2), \varphi_D(s_1+s_2), \omega\big(\varphi_D(s_1+s_2)\big)\Big) \ .$$

It shows that the group has the structure of semi-direct product of the multiplicative $\mathbb{R}_+$ with the additive $\mathbb{R}$ abelian groups ($t \in \mathbb{R}$, $e^s \in \mathbb{R}_+$). This group is usually called *affine group* $\mathcal{A}$ on $\mathbb{R}$. The time shifts form a normal subgroup of $\mathcal{A}$: they play the role of the abelian factor in the semi-direct product.

Time Reversal. The two solutions $\Psi_\eta(M, w; U, V)$ and $\Psi_{-\eta}(M, -w; U, V)$ are isometric. The isometry is implemented by the diffeomorphism φ_I defined by

$$\varphi_I : (U, V) \mapsto (-V, -U) \ . \tag{63}$$

It is well-defined everywhere in $\mathcal{M}$ and independent of η, M and w. Its action on the boundary is

$$U_\infty \mapsto -V_\infty \ , \quad V_\infty \mapsto -U_\infty \ , \quad T_0 \mapsto -T_0 \ , \tag{64}$$

and that on the physical phase space is

$$I : (\eta, M, w) \mapsto (-\eta, M, -w) \ . \tag{65}$$

The full group of symmetries generated by all three kinds of maps is denoted by G_σ. The index σ is to remind that the action of the group on $\mathcal{M}$ has been obtained with the help of our solution list that represents the gauge fixing σ; a different gauge fixing would lead to different action on $\mathcal{M}$. However, the action of G_σ on the physical phase space as well as some aspects of its action on the boundary are gauge independent. So is the group structure of G_σ, which is that of the affine group $\mathcal{A}_I$ with inversion.

The action of G_σ on the boundary $\partial\mathcal{M}$ defines a group of boundary transformations that will be denoted by G_b. The group G_b consists of pure diffeomorphisms (not polluted by any CCD's). Each element of G_σ determines a unique element of G_b and this map between G_σ and G_b is an isomorphism because G_σ acts truly on the boundary.

There is one point about the structure of G_b that will be useful later: the subgroup generated by time shifts and the time reversal is invariant (normal). Hence, there is a homeomorphism π_D of G_b onto the factor group, which is $G_{D\varphi}$, $\pi_D : G_b \mapsto G_{D\varphi}$. This means that the amount of dilatation hidden in any element of G_b is well-defined.

Decomposition of Diff$\mathcal{M}$. Let us choose a subgroup G of Diff$\mathcal{M}$ according to the following rules. The diffeomorphism φ is an element of G if

1. the map φ has a differentiable extension to the boundary of $\mathcal{M}$. Hence, it defines a map $\varphi_b : \partial\mathcal{M} \mapsto \partial\mathcal{M}$,
2. The action of φ on the boundary coincides with that of an element of the symmetry group G_σ, i.e., $\varphi_b \in G_b$.

The group G is a proper subgroup of Diff$\mathcal{M}$. It is the only part of Diff$\mathcal{M}$ that is interesting for the physics of the system. It is easy to check that the map $\pi_b : G \mapsto G_b$, the existence of which follows from the above definition, is a homeomorphism (preserves multiplication). The kernel G_0 of π_b is a set of elements of G that is mapped by π_b to the identity of G_b. Such a kernel must be a normal subgroup of G. The group G_0 is the *gauge group* of the system.

The group G itself needs a correction by CCD's at those elements that can be said to contain dilatation. The amount of dilatation in $\varphi \in G$ is obviously given by $\pi_D \pi_b \varphi$. Hence, the pairs $\big(\varphi, \omega(\pi_D \pi_b \varphi)\big)$ are *symmetries* of our system. They form a group with the same structure as G because CCD's commute with diffeomorphisms. Let us denote the group by G_c (c for corrected).

The action of the symmetry group G_c on the boundary has a gauge-invariant aspect. This can be revealed after attaching the scri to the solutions. Consider, for example, the spacetime $\big(\mathcal{M}, \Psi_+(M, w; U, V)\big)$. The $\mathcal{I}^+(+1, M, w)$ of this spacetime is one-dimensional and a coordinate $\tilde{U}$ can be chosen along it so that the difference $\tilde{U}_2 - \tilde{U}_1$ gives the interval of retarded time along $\mathcal{I}^+(+1, M, w)$. In the spherically symmetric case, such an interval is uniquely determined. However, $\tilde{U}$ clearly contains more information than just that about intervals: an origin of $\tilde{U}$ has also been chosen. This origin is the only non-trivial remainder, in the spherically symmetric case, of what can be called *asymptotic frame*. In an analogous way, an inertial coordinate system in Minkowski spacetime contains not only information about Minkowski interval (metric), but also that on the underlying inertial frame.

Let us attach $\mathcal{I}^+(+1, M, w)$ so that

$$\tilde{U} = U_\infty \ . \tag{66}$$

This is, in any case, possible because U_∞ *is* a retarded time along $\partial_+\mathcal{M}$ for the spacetime $\big(\mathcal{M}, \Psi_+(M, w; U, V)\big)$ so that the intervals match. However, (66) represents also a judicious choice of origins, that is, asymptotic frames, for each solution $\Psi_+(M, w; U, V)$. The idea behind the choice is that w keeps the role of the retarded arrival time of the out-going shells also with respect to $\mathcal{I}^+(+1, M, w)$. We shall see later how this choice of frame of asymptotic reference influences the form of the asymptotic action of the symmetry group and the calculation of the symplectic form.

At $\partial_-\mathcal{M}$, where $U \to -\infty$ and $V \in (u + 4M, \infty)$, we must recover the Schwarzschild advanced time $\tilde{V}$ as a function of V in order to attach $\mathcal{I}^-(+1, M, w)$. An easy calculation yields:

$$\tilde{V} = V_\infty + 4M \ln\left(\frac{-w + V_\infty}{4M} - 1\right) \ , \quad V_\infty = w + 4M\kappa\left[\exp\left(\frac{\tilde{V}}{4M}\right)\right] \ . \tag{67}$$

As $V \in (w + 4M, \infty)$, $\tilde{V}$ runs through $\mathbb{R}$. The advanced time interval is just $\tilde{V}_2 - \tilde{V}_1$. Equation (67) replaces (66) and determines the attachment of $\mathcal{I}^-(+1, M, w)$ to the spacetime $\big(\mathcal{M}, \Psi_+(M, w; U, V)\big)$. Observe that $\mathcal{I}^-(+1, M, w)$ takes only a part of $\partial_-\mathcal{M}$.

Finally, $\mathcal{I}^0(+1, M, w)$ is a "stretched" space-like infinity i^0: it has the structure of $\mathbb{R}$ and the coordinate T_∞ on it can be defined as limit point of a constant Schwarzschild time $T = T_\infty$ surfaces. Such surfaces are defined in the coordinates U and V by

$$\frac{U + \tilde{V}(V)}{2} = T_\infty \ . \tag{68}$$

We obtain from (67):

$$\frac{U}{2} + \frac{V}{2} + 2M \ln\left(\frac{-w+V}{4M} - 1\right) = \text{const} \tag{69}$$

as $U \to -\infty$ and $V \to \infty$. The surfaces are defined by the functions $\big(U(R), V(R)\big)$:

$$U = -R - 2M \ln\left(\frac{R}{2M} - 1\right) + T_\infty \tag{70}$$

and

$$V = w + 4M\kappa\left[\exp\left(\frac{T_\infty + R}{4M}\right)\sqrt{\frac{R}{2M} - 1}\right] . \tag{71}$$

The Schwarzschild time interval is the difference $T_{\infty 2} - T_{\infty 1}$. Again (67)–(71) not only imply that the intervals as calculated from $\tilde{V}$ or T_∞ match with intervals calculated from the geometry of $\big(\mathcal{M}, \Psi_+(M, w; U, V)\big)$. They also represent choices of origins, i.e., asymptotic frames, at $\mathcal{I}^-(+1, M, w)$ and $\mathcal{I}^0(+1, M, w)$, for each solution $\Psi_+(M, w; U, V)$.

The relation between the origins of the three spaces $\mathcal{I}^+$, $\mathcal{I}^-$ and $\mathcal{I}^0$ corresponding to *one* Schwarzschild spacetime (i.e., for a fixed M) can be set by a geometric convention. For example, the three surfaces $\tilde{U} = 0$, $\tilde{V} = 0$ and $T_\infty = 0$ (if the coordinates $\tilde{U}$, $\tilde{V}$ and T_∞ are extended into the spacetime as null or maximal surfaces) can be required to intersect all at a sphere of radius $R = 2M\kappa(1)$. This is the convention we use for each η, M and w. In this way, the retarded arrival time w can also be calculated by a standardized way from the description of the dynamics of the shell in an extended frame associated with $\mathcal{I}^-$ and $\mathcal{I}^0$. This is important for the action of time shifts as well as for the calculation of the Liouville form. The dependence of the geometric convention on the parameters η and M influences the form of asymptotic action of dilatations and the time reversal.

The attachment of $\mathcal{I}^+(-1, M, w)$, $\mathcal{I}^-(-1, M, w)$ and $\mathcal{I}^0(-1, M, w)$ to the spacetimes $\big(\mathcal{M}, \Psi_-(M, w; U, V)\big)$ is entirely analogous and can be skipped.

We observe that the scries have not, in general, a fixed position in the background manifold $\mathcal{M}$. This is due to the gauge fixing σ. Any gauge fixing implies also a definition of points (events) of $\mathcal{M}$ by some of their geometrical properties. For example, at the point (U, V), two light-like surfaces meet that have a particular geometrical meaning. The point at scri can also be defined by some geometrical properties. The latter definition does not coincides with a limit of the former in our case. This does not lead to any problems, if one does not try to use the definition of points inside $\mathcal{M}$ in an improper way: it is not gauge invariant and has not much physical meaning.

On the other hand, we can use the gauge fixing to calculate a lot of useful, gauge invariant results. For example, we can determine how the group G_c acts on the scri. Clearly, G_c acts there only via G_b so that, for example, G_0 must act trivially, leaving all points of scri invariant. For general $\varphi \in G$, a simple calculation reveals the following pattern. Let the diffeomorphism φ acts on the physical phase space as follows

$$\varphi : (\eta, M, w) \mapsto (\eta, M', w') \ , \tag{72}$$

that is, φ does not contain any time reversal. Then

$$\varphi_b : \mathcal{I}^+(\eta, M, w) \mapsto \mathcal{I}^+(\eta, M', w') \tag{73}$$

and the map is a *bijection* between the two scries. The same claim holds for $\mathcal{I}^-$ and $\mathcal{I}^0$. For example, the scri $\mathcal{I}^-(+1, M, w)$ has a past end point at $\partial_-\mathcal{M}$ with the value V_{past} of the coordinate V, $V_{\text{past}} = w + 4M$. The time shift $\varphi_H(t)$ sends V to $V + t$, hence $V_{\text{past}} \mapsto V_{\text{past}} + t$, and this is the past end point of $\mathcal{I}^-(+1, M, w + t)$.

Similarly, the time reversal

$$I : \mathcal{I}^+(\eta, M, w) \mapsto \mathcal{I}^-(-\eta, M, -w) \ , \tag{74}$$

etc., defines a bijection between the two scries.

Actually, we are going to identify all $\mathcal{I}^+$'s considering the points with the same value of the coordinate $\tilde{U}$ as equal. This gives a common $\mathcal{I}^+$ for all solutions. Similarly, common $\mathcal{I}^-$ and $\mathcal{I}^0$ can be defined. This structure is very important for the interpretation of the theory. We assume that observers are living at this common scri and are using the common frames (origins of time).

The definition of the common scri given above, together with the equations such as (73) and (74), imply that the group G_c acts on the common scri. This action is not difficult to calculate. The result is:

$$\begin{aligned} \varphi_H(t) &: (\tilde{U}, \tilde{V}, T_\infty) \mapsto (\tilde{U} + t, \tilde{V} + t, T_\infty + t) \ , \\ \varphi_D(s) &: (\tilde{U}, \tilde{V}, T_\infty) \mapsto (\tilde{U}e^s, \tilde{V}e^s, T_\infty e^s) \ , \\ \varphi_H(t) &: (\tilde{U}, \tilde{V}, T_\infty) \mapsto (-\tilde{V}, -\tilde{U}, -T_\infty) \, . \end{aligned} \tag{75}$$

The fact that the action of the time shift by t shifts all coordinates $\tilde{U}$, $\tilde{V}$ and T_∞ by the same amount is the consequence of the judicious choice of origins. Similarly, the homogeneous action of the dilatations is due to correlations of our choices of frames for all different values of the parameter M.

Clearly, the group G_c acts transitively on Γ. If the symplectic structure of the physical phase space Γ were known, the functions generating the infinitesimal transformations of the group could be found. Then the dynamics based on the time shifts could be constructed.

3.2 Canonical Theory

The variables η, M and w can play the role of Dirac observables, but we do not know what are their Poisson brackets. In the present section, we calculate the brackets. We start from a Hamiltonian action principle for null shells and reduce it to physical phase space applying the technique invented by Kuchař.

The Action. As a Hamiltonian action principle that implies the dynamics of our system, we take the action Eq. (2.6) of [30] (see also [31]). Let us briefly summarize the relevant formulae. The spherically symmetric metric is written in the form:

$$ds^2 = -N^2 d\tau^2 + \Lambda^2 (d\rho + N^\rho d\tau)^2 + R^2 d\Omega^2 \ , \tag{76}$$

the shell is described by its radial coordinate $\rho = \mathbf{r}$ and its conjugate momentum $\mathbf{p}$. The action reads

$$S_0 = \int d\tau \left[\mathbf{p}\dot{\mathbf{r}} + \int d\rho (P_\Lambda \dot{\Lambda} + P_R \dot{R} - H_0) \right] \ , \tag{77}$$

and the Hamiltonian is

$$H_0 = N\mathcal{H} + N^\rho \mathcal{H}_\rho \ + N_\infty E_\infty \ , \tag{78}$$

where N and N^ρ are the lapse and shift functions, $\mathcal{H}$ and $\mathcal{H}_\rho$ are the constraints,

$$\begin{aligned} \mathcal{H} \ &= \mathrm{G} \left(\frac{\Lambda P_\Lambda^2}{2R^2} - \frac{P_\Lambda P_R}{R} \right) + \frac{1}{\mathrm{G}} \left(\frac{RR''}{\Lambda} - \frac{RR'\Lambda'}{\Lambda^2} + \frac{R'^2}{2\Lambda} - \frac{\Lambda}{2} \right) \\ &\quad + \frac{\eta \mathbf{p}}{\Lambda} \delta(\rho - \mathbf{r}) \ , &(79) \\ \mathcal{H}_\rho &= P_R R' - P'_\Lambda \Lambda - \mathbf{p}\delta(\rho - \mathbf{r}) \ , &(80) \end{aligned}$$

and the prime or dot denote the derivatives with respect to ρ or τ. The term $N_\infty E_\infty$ is the remainder of $\mathcal{H}_\infty[\boldsymbol{N}_\infty]$ (see Sect. 2.2) in the case of rotational symmetry. $N_\infty := \lim_{\rho \to \infty} N(\rho)$ and E_∞ is the ADM mass (see [30]). In the Schwarzschild spacetime with mass parameter M and asymptotic Schwarzschild time T_∞, it holds that

$$N_\infty E_\infty = \frac{1}{\mathrm{G}} M \dot{T}_\infty \ . \tag{81}$$

The term can then be transfered from the Hamiltonian to the part of action containing time derivatives – the so-called *Liouville form*, see [32].

The Hamiltonian constraint function $\mathcal{H}$ is written in a way that differs from [30] and [6] in that its dependence on the Newton constant G becomes visible. This is necessary in order that the scaling properties[6] of $\mathcal{H}$ are manifest, $\mathcal{H} \mapsto e^{-s}\mathcal{H}$. Observe that the rescaling of $\mathcal{H}$ cannot be implemented by a canonical transformation: all function that might be tried to generate such rescaling had vanishing Poisson brackets with G. This does not mean, however, that vacuum Einstein equations are not invariant under dilatation (cf. [33]). In our case, the geometric sector alone (see Sect. 3.1) as well as the matter sector alone, are invariant under dilatation, but the whole theory is not: the relation between geometry and matter involves the Newton constant.

---

[6] To reveal the scaling behaviour, units must be chosen so that $\hbar = c = 1$, but $\mathrm{G} \neq 1$!

The "volume" variables Λ, R, P_Λ, P_R, N and N^ρ are the same as in [30] and [6]. The meaning of the variables Λ, R, N and N^ρ can be inferred from the spacetime metric (76). The momenta conjugate to the configuration variables Λ and R can be calculated from the action S_0 by varying it with respect to P_Λ and P_R:

$$P_\Lambda = -\frac{R}{\mathrm{G}N}(\dot{R} - N^\rho R') \;, \tag{82}$$

and

$$P_R = -\frac{\Lambda}{\mathrm{G}N}(\dot{R} - N^\rho R') - \frac{R}{\mathrm{G}N}[\dot{\Lambda} - (N^\rho \Lambda)'] \;. \tag{83}$$

Some differentiability conditions at the shell are important in order that proper equations of motion are obtained by varying the action. One can assume as in [30] that the gravitational variables are smooth functions of ρ, with the exception that N', $(N^\rho)'$, Λ', R', P_Λ and P_R may have finite discontinuities at isolated values of ρ. The coordinate loci of the discontinuities are smooth functions of τ for each shell. This follows from (i) the conditions at shell points which are the same as in [30] and (ii) from the corresponding choice of foliation: the metric with respect to coordinates τ and ρ may be piecewise smooth and everywhere continuous.

The Liouville Form at the Constraint Surface. Our aim is to calculate the Poisson brackets between Dirac observables such as M and w. We can employ the property of the pull-back $\Theta_\mathcal{C}$ of the Liouville form Θ to the constraint surface $\mathcal{C}$ that it depends only on the Dirac observables as it has been explained in Sect. 2 (see also [9], [3] and [4]). Its external differential then defines the symplectic form of the physical phase space, which determine the brackets. Observe further that the pull-back $\Theta_\mathcal{C}$, if integrated over τ, gives the action of the reduced system. Indeed, if one solves the constraints, only the Liouville form remains from the action S_0.

Thus, we have to transform the Liouville form to the variables M, w and a set of observable-dependent embeddings; these variables form a coordinate system on the constraint surface for each case $\eta = +1$ and $\eta = -1$.

An important point is to specify the family of embeddings that will be used. The embeddings are given by

$$U = U(o, \rho), \quad V = V(o, \rho) \;, \tag{84}$$

where U and V are the coordinates defined in Sect. 2.1. These functions have to satisfy several conditions.

1. As Σ is space-like, U and V are null and increasing towards the future, we must have $U' < 0$ and $V' > 0$ everywhere.
2. At the regular centre, the four-metric is flat and the three-metric is to be smooth. This implies, for all o's, that $U'(o, 0) = -V'(o, 0)$ in addition to

the condition $U(o,0) = V(o,0)$. This follows from $T'(U(o,0),V(o,0)) = 0$ and means that Σ must run parallel to $T = \text{const}$ in order to avoid conical singularities. Here, $T = (U+V)/2$ is the time of the inertial system at the centre.

3. At the space-like infinity, the four-metric is the Schwarzschild metric. We require that the embedding approaches the Schwarzschild-time-constant surfaces $T = \text{const}$, and that ρ becomes the Schwarzschild curvature coordinate R asymptotically. More precisely, the behaviour of the Schwarzschild coordinates T and R along each embedding $U(o,\rho), V(o,\rho)$ must satisfy
$$T(\rho) = T_\infty + O(\rho^{-1}), \tag{85}$$
$$R(\rho) = \rho + O(\rho^{-1}) \; . \tag{86}$$
The asymptotic coordinate T_∞ is a gauge-invariant quantity and it possesses the status of an observable. It follows then from (70) and (71) that the embeddings must depend on η, M and w.
4. At the shell ($\rho = \mathbf{r}$) we require $U(o,\rho)$ and $V(o,\rho)$ to be C^∞ functions of ρ. In fact, as the four-metric is continuous in the coordinates U and V, but not smooth, only the C^1-part of this condition is gauge invariant. Jumps in all higher derivatives are gauge dependent, but the condition will simplify equations without influencing results.

The Liouville form of the action (37) can be written as follows:
$$\Theta = \mathbf{p}\dot{\mathbf{r}} - \frac{1}{\mathrm{G}} M\dot{T}_\infty + \int_0^\infty d\rho\,(P_\Lambda\dot{\Lambda} + P_R\dot{R})\,. \tag{87}$$

We can now start to transform (87) into Kuchař variables. It is advantageous to let first the double-null coordinates U and V arbitrary and the Dirac observables o^i, $i = 1,\ldots,2n$ unspecified. We just need to know that the metric and the embeddings depends on o^i:
$$A = A(U,V;o)\ , \quad R = R(U,V;o)\ , \quad U = U(o,\rho)\ , \quad V = V(o,\rho)\ . \tag{88}$$
For any double-null gauge, the equations
$$4RR_{,UV} + 4R_{,U}R_{,V} + A = 0\ , \tag{89}$$
$$AR_{,UU} - A_{,U}R_{,U} = 0\ , \tag{90}$$
$$AR_{,VV} - A_{,V}R_{,V} = 0 \tag{91}$$
represent the condition that the transformation is performed at the constraint surface Γ (cf. [6], Eqs. (32)–(34)).

The transformation formulae for Λ, R, N and N^ρ can be read off the metric:
$$R = R\big(U(o,\rho),V(o,\rho),o\big)\ , \tag{92}$$
$$\Lambda = \sqrt{-AU'V'}\ , \tag{93}$$
$$\mathcal{N} = -\frac{\dot{U}V' - \dot{V}U'}{2U'V'}\sqrt{-AU'V'}\ , \tag{94}$$
$$\mathcal{N}^\rho = \frac{\dot{U}V' + \dot{V}U'}{2U'V'}\ . \tag{95}$$

The transformation of the momenta P_R and P_Λ are obtained from their definitions (82) and (83), into which (92)–(95) are substituted:

$$\mathrm{G}P_\Lambda = \frac{R}{\sqrt{-AU'V'}}\,(R_{,U}U' - R_{,V}V')\,, \tag{96}$$

$$\mathrm{G}P_R = R_{,U}U' - R_{,V}V' + \frac{RA_{,U}}{2A}U' - \frac{RA_{,V}}{2A}V' + \frac{R}{2}\frac{U''}{U'} - \frac{R}{2}\frac{V''}{V'}\,. \tag{97}$$

The transformation of the shell variable $\mathbf{r}$ follows from the obvious relations

$$U(o,\mathbf{r}) = w\ , \quad V(o,\mathbf{r}) = w\ , \tag{98}$$

each valid in one of the cases $\eta = +1$ or $\eta = -1$. Finally, the value of the momentum $\mathbf{p}$ at the constraint surface can be found if the differentiability properties of the volume functions at the shell are used and the coefficients at the δ-functions in the constraints (79) and (80) are compared (for details see [6]):

$$\mathbf{p} = -\eta R\Delta_{\mathbf{r}}(R') = -\mathrm{G}\Lambda\Delta_{\mathbf{r}}(P_\Lambda)\ , \tag{99}$$

where $\Delta_{\mathbf{r}}(f)$ denotes a jump in the function f across the shell at the point $\rho = \mathbf{r}$.

The form (87) can be divided into a boundary part (the first two terms on the right-hand side) and the volume parts. Each volume part is associated with a particular component of the space cut out by the shell; it has the form

$$\int_a^b d\rho\,(P_\Lambda\dot\Lambda + P_R\dot R)\ , \tag{100}$$

where a and b are values of the coordinate ρ at the boundary of the volume. There are only two cases, $a = 0$, $b = \mathbf{r}$ and $a = \mathbf{r}$, $b = \infty$.

Since the pull-back of the Liouville form cannot depend on the volume variables $U(o,\rho)$ and $V(o,\rho)$, its volume part must be, after the transformation, given by

$$\Theta_a^b|_{\mathcal{C}} = \int_a^b d\rho\,[(f\dot U + g\dot V + h_i\dot o^i)' + \dot\varphi]\ , \tag{101}$$

cf. [6]. Comparing the coefficients at the dotted and primed quantities in (100) and (101) leads to partial differential equations for the functions f, g and h_i [6] and one can show that (89)–(91) are the integrability conditions for these partial differential equations.

In such a way, $\Theta_{\mathcal{C}}$ can be transformed to a sum of boundary terms. These terms can be calculated and simplified if the transformation equations for the shell variables and the properties of the embeddings at the boundary points are used. It is a lengthy and rather technical calculation that is not very interesting for us now; it has been described in detail in [6] and, especially, [28]. We skip it and quote only the final results:

$$\eta = +1: \quad \Theta|_{\mathcal{C}} = -\frac{1}{\mathrm{G}}M\dot w\ ; \quad \eta = -1: \quad \Theta|_{\mathcal{C}} = -\frac{1}{\mathrm{G}}M\dot w\ . \tag{102}$$

Only one aspect of the calculation deserves mentioning. This is the cancellation of the term $-(M/\mathrm{G})\dot{T}_\infty$ in the Liouville form (87) by a term coming from the volume part (100). The cancellation is again a result of our judicious choice of asymptotic frames as described in Sect. 3.1: the variable T_∞ enters the calculation via (70) and (71) that are implied by the choice.

At this stage, it is advantageous to introduce new variables for the description of the Dirac observables, that is, for the matter sector in our case: let us define, for $\eta = +1$:

$$p_u := -\frac{M}{\mathrm{G}} \ , \quad u := w \ , \tag{103}$$

and for $\eta = -1$,

$$p_v := -\frac{M}{\mathrm{G}} \ , \quad v := w \ . \tag{104}$$

Observe that the momenta p_u and p_v have the meaning of negative energy and scale in accord with this meaning as e^{-s}. The rescaling of p's, together with the rescaling by e^s of u and v, can be implemented canonically.

The desired Poisson brackets can be written down immediately:

$$\{u, p_u\} = 1 \ , \quad \{v, p_v\} = 1 \ . \tag{105}$$

the generator of the infinitesimal time shift is $p_u dt$ or $p_v dt$ and that of dilatation is $p_u u ds$ or $p_v v ds$.

In [28], a general method of integration of the differential equations for the functions f, g and h_i has been developed and a formula generalizing (102) to a system containing any number of in-going and any number out-going shell and the due shell intersections (Kouletsis formula) has been written down.

The reduced action that contains only the physical degrees of freedom is, if $\eta = +1$,

$$S_+ = \int d\tau p_u \dot{u} \ , \tag{106}$$

and if $\eta = -1$,

$$S_- = \int d\tau p_v \dot{v} \ . \tag{107}$$

These actions imply the equations of motion

$$p_u = \mathrm{const} \ , \quad v = \mathrm{const} \tag{108}$$

or

$$p_v = \mathrm{const} \ , \quad v = \mathrm{const} \ , \tag{109}$$

as they are to be.

Merging In- and Out-Going Dynamics. In this subsection, the two reduced action principles for the out-going and in-going dynamics of the shell will be recognized as a result of reduction of a single merger action. The merger action gives a complete account of all motions of the shell on the background manifold $\mathcal{M}$. The merging is a step of certain significance for our quantum theory.

Consider the action (106) for out-going shells. The variables p_u and u do not contain the full information about where the shell is in $\mathcal{M}$. The function $u(\tau)$ is the value of only one coordinate U of the shell for the value τ of the time parameter. For a complete description, $v(\tau)$ would have to be added. For a solution trajectory, $u(\tau)$ is constant, but $v(\tau)$ can be arbitrary: it depends on the choice of the parameter τ and so it can be considered as a gauge variable in a reparametrization-invariant formalism. Hence, a valid extension of the action (106) is

$$\bar{S}_+ = \int d\tau (p_u \dot{u} + p_v \dot{v} - n_+ p_v) \ , \tag{110}$$

where n_+ is a Lagrange multiplier that enforces the vanishing of the momentum p_v conjugate to the gauge variable v. The meaning of n_+ transpires from the equation

$$\dot{v} = n_+ \tag{111}$$

obtained by varying $\bar{S}_+$ with respect to p_v. Thus, n_+ measures the rate of τ with respect to the parameter $2t$: for any out-going shell, we have

$$t = \frac{1}{2}\big(u + v(\tau)\big) \ , \tag{112}$$

where t is the time of the inertial frame defined by the regular centre $R = 0$ and, at the same time, it coincides with the parameter of the time-shift group.

We can define another Lagrangian multiplier n by

$$n_+ = n p_u \ , \tag{113}$$

because p_u is always non-zero in the case $\eta = +1$. The multiplier n measures the rate of τ with respect to the "physical parameter": the corresponding tangent vector to particle trajectory coincides with the contravariant four-momentum of the particle.

An analogous extension $\bar{S}_-$ of S_-, (107) is

$$\bar{S}_- = \int d\tau (p_u \dot{u} + p_v \dot{v} - n_- p_u) \ , \tag{114}$$

and we can switch to n defined by $n_- = n p_v$ now.

The merger action is

$$\bar{S} = \int d\tau (p_u \dot{u} + p_v \dot{v} - n p_u p_v) \ . \tag{115}$$

It contains both cases as the two possible solutions of the constraint

$$p_u p_v = 0 \ : \tag{116}$$

if $p_v = 0$, we obtain the case $\eta = +1$, and if $p_u = 0$, $\eta = -1$.

The two variables u and v determine the position of the shell in $\mathcal{M}$ uniquely. The generator of the time translation is $p_u + p_v$, the value of which is the negative of the total energy $E = M/\mathrm{G}$ of the system, and the generator of the dilatation is $p_u u + p_v v$.

The new phase space has non-trivial boundaries:

$$p_u \leq 0, \quad p_v \leq 0 \ , \tag{117}$$

$$\frac{-u+v}{2} > 0 \ , \tag{118}$$

$$p_v = 0 \ , \quad U \in (-\infty, u) \ , \quad V > u - 4p_u\kappa\left(-\exp\frac{u-U}{4p_u}\right) \ , \tag{119}$$

and

$$p_u = 0 \ , \quad V \in (v, \infty) \ , \quad U < v + 4p_v\kappa\left(-\exp\frac{V-v}{4p_v}\right) \ . \tag{120}$$

The boundaries defined by inequalities (119) and (120) are due to the singularity. Equations (119) and (120) limit the values of U and V that can be used for embeddings.

The two dynamical systems defined by the actions (77) and (115) are equivalent: each maximal dynamical trajectory of the first, if transformed to the new variables, gives a maximal dynamical trajectory of the second and vice versa.

The variables u, v, p_u and p_v span the effective phase space of the shell. They contain all true degrees of freedom of the system. One can observe that the action (115) coincides with the action for free motion of a zero-rest-mass spherically symmetric (light-like) shell in flat spacetime. Such a dynamics is complete because there is no geometric singularity at the value zero of the radius of the shell, $(-u+v)/2$, and this point can be considered as a harmless caustic so that the light can re-expand after passing through it. The dynamics of the physical degrees of freedom by itself is, therefore, regular.

It might seem possible to extend the phase space of the gravitating shell, too, in the same way so that the in-going and the out-going sectors are merged together into one bouncing solution. However, such a formal extension of the dynamics (115) is not adequate. The physical meaning of any solution written in terms of new variables is given by measurable quantities of geometrical or physical nature, which include now also the curvature of spacetime. These observables must be expressed as functions on the phase space. They can of course be transformed between the phase spaces of the two systems (77) and (115).

They cannot be left out from any complete description of a system, though they are often included only tacitly: an action alone does not define a system. This holds just as well for the action (77) as for (115).

Let us consider these observables. Equations (31)–(34) and (39)–(42) can be used to show that the curvature of the solution spacetime diverges at the boundary defined by (119) for $p_v = 0$ and by (120) for $p_u = 0$. It follows that the observable quantities at and near the "caustic" are badly singular and that there is no sensible extension of the dynamics defined by action (115) to it, let alone through it. This confirms the more or less obvious fact that no measurable property (such as the singularity) can be changed by a transformation of variables.

The action for the null dust shell is now written in a form which can be taken as the starting point for quantization. Surprisingly, it will turn out that a quantum theory can be constructed so that it is singularity-free. This will be shown in the next section.

4 Quantum Theory

In this section, we shall construct a quantum theory of our model. Of course, there is no unique construction of a quantum theory for any given classical model. We just show that there is a singularity-free, unitary quantum mechanics of the shell. The account in this section follows [34] adding a few new ideas.

4.1 Group Quantization

To quantize the system defined by the action (115), we apply the so-called group-theoretical quantization method [35]. There are three reasons for this choice. First, the method as modified for the generally covariant systems by Rovelli [36] (see also [37] and [38]) is based on the algebra of Dirac observables of the system; dependent degrees of freedom don't influence the definition of Hilbert space. Second, the group method has, in fact, been invented to cope with restrictions such as (117) and (118). By and large, one has to choose a set of observables that form a Lie algebra; this algebra has to generate a group of symplectomorphisms that has to act transitively in the phase space respecting all boundaries. In this way, the information about the boundaries is built in the quantum mechanics. Finally, the method automatically leads to self-adjoint operators representing all observables. In particular, a self-adjoint extension of the Hamiltonian is obtained in this way, and this is the reason that the dynamics is unitary.

To begin with, we have to find a complete set of Dirac observables. Let us choose the functions p_u, p_v, $D_u := up_u$ and $D_v := vp_v$. Observe that u alone is constant only along out-going shell trajectories ($p_u \neq 0$), and v only along in-going ones ($p_v \neq 0$), but up_u and vp_v are always constant. The only non vanishing Poisson brackets are

$$\{D_u, p_u\} = p_u, \quad \{D_v, p_v\} = p_v. \tag{121}$$

This Lie algebra generates a group Q_2 of symplectic transformations of the phase space that preserve the boundaries $p_u = 0$ and $p_v = 0$. Q_2 is the Cartesian product of two copies of the affine group $\mathcal{A}$ on $\mathbb{R}$.

The group $\mathcal{A}$ generated by p_u and D_u has three irreducible unitary representations. In the first one, the spectrum of the operator $\hat{p}_u$ is $[0, \infty)$, in the second, $\hat{p}_u$ is the zero operator, and in the third, the spectrum is $(-\infty, 0]$, see [39]. Thus, we must choose the third representation; this can be described as follows (details are given in [39]).

The Hilbert space is constructed from complex functions $\psi_u(p)$ of $p \in [0, \infty)$; the scalar product is defined by

$$(\psi_u, \phi_u) := \int_0^\infty \frac{dp}{p} \; \psi_u^*(p)\phi_u(p), \tag{122}$$

and the action of the generators $\hat{p}_u$ and $\hat{D}_u$ on smooth functions is

$$(\hat{p}_u\psi_u)(p) = -p\psi_u(p), \quad (\hat{D}_u\psi_u)(p) = -ip\frac{d\psi_u(p)}{dp}. \tag{123}$$

Similarly, the group generated by p_v and D_v is represented on functions $\psi_v(p)$; the group Q_2 can, therefore, be represented on pairs $\Big(\psi_u(p), \psi_v(p)\Big)$ of functions:

$$\hat{p}_u\Big(\psi_u(p), \psi_v(p)\Big) = \Big(-p\psi_u(p), 0\Big), \tag{124}$$

$$\hat{p}_v\Big(\psi_u(p), \psi_v(p)\Big) = \Big(0, -p\psi_v(p)\Big), \tag{125}$$

$$\hat{D}_u\Big(\psi_u(p), \psi_v(p)\Big) = \Big(-ip\frac{d\psi_u(p)}{dp}, 0\Big), \tag{126}$$

$$\hat{D}_v\Big(\psi_u(p), \psi_v(p)\Big) = \Big(0, -ip\frac{d\psi_v(p)}{dp}\Big). \tag{127}$$

This choice guarantees that the Casimir operator $\hat{p}_u\hat{p}_v$ is the zero operator on this Hilbert space, and so the constraint is satisfied. The scalar product is

$$\Big(\big(\psi_u(p), \psi_v(p)\big), \big(\phi_u(p), \phi_v(p)\big)\Big) = \int_0^\infty \frac{dp}{p}\big(\psi_u^*(p)\phi_u(p) + \psi_v^*(p)\phi_v(p)\big) . \tag{128}$$

Let us call this representation $\mathcal{R}_2$. Observe that this representation of the group $\mathcal{A} \times \mathcal{A}$ is not irreducible. It is, however, irreducible for the group extended by the time reversal because its representative swaps the invariant subspaces. The representation $\mathcal{R}_2$ has a well-defined meaning as a quantum theory of our system. In it, the out-going shells are independent of the in-going ones. The dynamics of each kind of shells is complete for itself. One could study this dynamics by defining a position operator $\hat{r}(t)$ in a natural way similar to what will be done later and one would find that the in-going shells simply proceed through $r = 0$ into negative values of r. Analogous holds, in the time reversed order, for the out-going shells.

Handling the inequality (118) is facilitated by the canonical transformation:

$$t = (u + v)/2, \qquad r = (-u + v)/2, \tag{129}$$
$$p_t = p_u + p_v, \qquad p_r = -p_u + p_v. \tag{130}$$

The constraint function then becomes $p_u p_v = (p_t^2 - p_r^2)/4$.

The function $p_t \delta t$ generates, via Poisson brackets, the infinitesimal time shift in (t, r)-space, $t \mapsto t + \delta t$, $r \mapsto r$, and $p_r \delta r$ generates an r-shift, $t \mapsto t$, $r \mapsto r + \delta r$. We introduce also the observables

$$D := D_u + D_v = tp_t + rp_r \ , \tag{131}$$

and

$$J := -D_u + D_v = rp_t + tp_r \ . \tag{132}$$

The function $D\delta s$ generates a dilatation in the (t, r)-space, $t \mapsto t + t\delta s$, $r \mapsto r + r\delta s$, and $J\delta v$ generates a boost, $t \mapsto t + r\delta v$, $r \mapsto r + t\delta v$. What do these transformations with our half-plane $r > 0$? The transformations generated by p_t and D preserve the boundary while those by p_r and J do not. Hence, only the subgroup Q_1 of Q_2 generated by p_t and D respects the inequality (118).

The representation $\mathcal{R}_2$ of Q_2 can be decomposed into the direct sum of two equivalent representations of Q_1. This representation of Q_1 will be denoted by $\mathcal{R}_1$ and it will serve as our *definitive* quantum mechanics. Let us stress that this quantum mechanics does not describe a subsystem of the physical system described with the representation $\mathcal{R}_2$. The representation $\mathcal{R}_1$ can be described as follows: The states are determined by complex functions $\varphi(p)$ on $\mathbb{R}_+$; the scalar product (φ, ψ) is

$$(\varphi, \psi) = \int_0^\infty \frac{dp}{p} \ \varphi^*(p)\psi(p); \tag{133}$$

let us denote the corresponding Hilbert space by $\mathcal{K}$. The representatives of the above algebra are

$$(\hat{p}_t\varphi)(p) \ = -p\varphi(p), \tag{134}$$
$$(\hat{p}_r^2\varphi)(p) \ = p^2\varphi(p), \tag{135}$$
$$(\hat{D}\varphi)(p) \ = -ip\frac{d\varphi(p)}{dp}, \tag{136}$$
$$(\hat{J}^2\varphi)(p) = -p\frac{d\varphi(p)}{dp} - p^2\frac{d^2\varphi(p)}{dp^2}. \tag{137}$$

There are also well-defined operators for p_r^2 and J^2 because, in the representation $\mathcal{R}_2$, the identities $\hat{p}_r^2 = \hat{p}_t^2$ and $\hat{J}^2 = \hat{D}^2$ hold.

An important observation is that $\mathcal{R}_1$ is a representation of the group of asymptotic symmetries as described in Sect. 3.1. Another important observation is that the quantum mechanics $\mathcal{R}_2$ describes two discoupled degrees of freedom, namely the in- and out-going shells, while $\mathcal{R}_1$ describes a system with a single degree of freedom: the in- and out-going motions have been coupled into one motion. Let us study, what is this motion.

First, we have to construct a time evolution. For that, we use the time shift symmetry generated by $\hat{p}_t$. The operator $-\hat{p}_t$ has the meaning of the total energy $E = M/\mathrm{G}$ of the system. We observe that it is a self-adjoint operator with a positive spectrum and that it is diagonal in our representation. The parameter t of the unitary group $\hat{U}(t)$ that is generated by $-\hat{p}_t$ is easy to interpret: t represents the quantity that is conjugated to p_t in the classical theory and this is given by (129). Hence, $\hat{U}(t)$ describes the evolution of the shell states, for example, between the levels of the function $(U+V)/2$ on $\mathcal{M}$.

The missing piece of information of where the shell is on $\mathcal{M}$ is carried by the quantity r of (129). We shall define the corresponding position operator in three steps.

First, we observe that r itself is not a Dirac observable, but the boost J is, and that the value of J at the surface $t=0$ coincides with rp_t. It follows that the meaning of the Dirac observable Jp_t^{-1} is the position at the time $t=0$. This is in a nice correspondence with the Newton-Wigner construction [40] on one hand, and with the notion of evolving constants of motion by Rovelli [27] on the other.

Second, we try to make Jp_t^{-1} into a symmetric operator on our Hilbert space. However, in the representation $\mathcal{R}_1$, only $\hat{J}^2$ is meaningful. Let us then chose the following factor ordering for $J^2p_t^{-2}$:

$$\hat{r}^2 := \frac{1}{\sqrt{p}}\hat{J}\frac{1}{p}\hat{J}\frac{1}{\sqrt{p}} = -\sqrt{p}\frac{d^2}{dp^2}\frac{1}{\sqrt{p}}. \tag{138}$$

Other choices are possible; the above one makes $\hat{r}^2$ essentially a Laplacian and this simplifies the subsequent mathematics. Indeed, we can map $\mathcal{K}$ unitarily to $L^2(\mathbb{R}_+)$ by sending each function $\psi(p) \in \mathcal{K}$ to $\tilde{\psi}(p) \in L^2(\mathbb{R}_+)$ as follows:

$$\tilde{\psi}(p) = \frac{1}{\sqrt{p}}\psi(p). \tag{139}$$

Then, the operator of squared position $\tilde{r}^2$ on $L^2(\mathbb{R}_+)$ corresponding to $\hat{r}^2$ is

$$\tilde{r}^2\tilde{\psi}(p) = \frac{1}{\sqrt{p}}\hat{r}^2\left(\sqrt{p}\tilde{\psi}(p)\right) = -\frac{d^2\tilde{\psi(p)}}{dp^2} = -\tilde{\Delta}\tilde{\psi}(p). \tag{140}$$

Third, we have to extend the operator $\hat{r}^2$ to a self-adjoint one. The Laplacian on the half-axis possesses a one-dimensional family of such extensions [41]. The parameter is $\alpha \in [0,\pi)$ and the domain of $\tilde{\Delta}_\alpha$ is defined by the boundary condition at zero:

$$\tilde{\psi}(0)\sin\alpha + \tilde{\psi}'(0)\cos\alpha = 0. \tag{141}$$

The complete system of normalized eigenfunctions of $\tilde{\Delta}_\alpha$ can then easily be calculated:

$$\tilde{\psi}_\alpha(r,p) = \sqrt{\frac{2}{\pi}}\,\frac{r\cos\alpha\cos rp - \sin\alpha\sin rp}{\sqrt{r^2\cos^2\alpha + \sin^2\alpha}}; \tag{142}$$

if $\alpha \in (0, \pi/2)$, there is one additional bound state,

$$\tilde{\psi}_\alpha(b,p) = \frac{1}{\sqrt{2\tan\alpha}}\exp(-p\tan\alpha), \tag{143}$$

so that

$$-\tilde{\Delta}_\alpha\tilde{\psi}_\alpha(r,p) = r^2\tilde{\psi}_\alpha(r,p), \tag{144}$$
$$-\tilde{\Delta}_\alpha\tilde{\psi}_\alpha(b,p) = -\tan^2\alpha\,\tilde{\psi}_\alpha(r,p). \tag{145}$$

The corresponding eigenfunctions ψ_α of the operator $\hat{r}^2_\alpha$ are:

$$\psi_\alpha(r,p) = \sqrt{\frac{2p}{\pi}}\,\frac{r\cos\alpha\cos rp - \sin\alpha\sin rp}{\sqrt{r^2\cos^2\alpha + \sin^2\alpha}}, \tag{146}$$

and we restrict ourselves to $\alpha \in [\pi/2, \pi]$, so that there are no bound states and the operator $\hat{r}$ is self-adjoint. Indeed, $\tilde{r}$ is the square root of $-\tilde{\Delta}_\alpha$, hence its eigenvalue for the bound state is imaginary.

To restrict the choice further, we apply the idea of Newton and Wigner [40]. First, the subgroup of Q_0 that preserves the surface $t = 0$ is to be found. This is, in our case, $U_D(s)$ generated by the dilatation D. Then, in the quantum theory, the eigenfunctions of the position at $t = 0$ are to transform properly under this group; this means that the eigenfunction for the eigenvalue r is to be transformed to that for the eigenvalue $e^{-s}r$, for each s and r. The dilatation group generated by $\hat{D}$ acts on a wave function $\psi(p)$ as follows:

$$\psi(p) \mapsto U_D(s)\psi(p) = \psi(e^{-s}p), \tag{147}$$

where $U_D(s)$ is an element of the group parameterized by s. Applying $U_D(s)$ to $\psi_\alpha(r,p)$ yields

$$U_D(s)\psi_\alpha(r,p) = e^{-s/2}\sqrt{\frac{2p}{\pi}}\,\frac{r\cos\alpha\cos(e^{-s}rp) - \sin\alpha\sin(e^{-s}rp)}{\sqrt{r^2\cos^2\alpha + \sin^2\alpha}}. \tag{148}$$

The factor $e^{-s/2}$ in the resulting functions of p keeps the system δ-normalized.

Let $\alpha = \pi/2$; then

$$U_D(s)\psi_{\pi/2}(r,p) = e^{-s/2}\psi_{\pi/2}(e^{-s}r,p). \tag{149}$$

Similarly, for $\alpha = \pi$,

$$U_D(s)\psi_\pi(r,p) = e^{-s/2}\psi_\pi(e^{-s}r,p), \tag{150}$$

but such a relation can hold for no other α from the interval $[\pi/2, \pi]$, because of the form of the eigenfunction dependence on r. Now, Newton and Wigner require that

$$U_D(s)\psi(r,p) = e^{-s/2}\psi(e^{-s}r,p). \tag{151}$$

Then all values of α except for $\alpha = \pi/2$ and $\alpha = \pi$ are excluded.

We have, therefore, only two choices for the self-adjoint extension of $\hat{r}^2$:

$$\psi(r,p) := \sqrt{\frac{2p}{\pi}}\sin rp, \quad r \geq 0, \tag{152}$$

and

$$\psi(r,p) := \sqrt{\frac{2p}{\pi}}\cos rp, \quad r \geq 0. \tag{153}$$

Let us select the first set, (152); by that, the construction of a position operator is finished. The construction contains a lot of choice: the large factor-ordering freedom, and the freedom of choosing the self-adjoint extension.

Another observable that we shall need is $\hat{\eta}$; this is to tell us the direction of motion of the shell at the time zero, having the eigenvalues $+1$ for all purely out-going shell states, and -1 for the in-going ones. In fact, in the classical theory, $\eta = -\mathrm{sgn}p_r$, but p_r does not act as an operator on the Hilbert space $\mathcal{K}$, only p_r^2, and the sign is lost.

Consider the classical dilatation generator $D = tp_t + rp_r$. It is a Dirac observable; at $t = 0$, its value is rp_r. Thus, for positive r, the sign of $-D$ at $t = 0$ has the required value. Hence, we have the relation:

$$\mathrm{sgn}D = -\eta_{t=0}. \tag{154}$$

The eigenfunctions $\psi_a(p)$ of the operator $\hat{D}$ are solutions of the differential equation:

$$\hat{D}\psi_a(p) = a\psi_a(p). \tag{155}$$

The corresponding normalized system is given by

$$\psi_a(p) = \frac{1}{\sqrt{2\pi}}p^{ia}. \tag{156}$$

Hence, the kernels $P_\pm(p,p')$ of the projectors $\hat{P}_\pm$ on the purely out- or in-going states are:

$$P_+(p,p') = \int_{-\infty}^0 da\,\psi_a(p)\frac{\psi_a^*(p')}{p'}, \quad P_-(p,p') = \int_0^\infty da\,\psi_a(p)\frac{\psi_a^*(p')}{p'} \tag{157}$$

so that

$$(\hat{\eta}_0\psi)(p) = \int_0^\infty dp'[P_+(p,p') - P_-(p,p')]\psi(p'). \tag{158}$$

The observables $\hat{p}_t$, $\hat{D}$, $\hat{r}_0$ and $\hat{\eta}_0$ will suffice to work out a number of interesting predictions.

4.2 Motion of Wave Packets

We shall work with the family of wave packets on the energy half-axis that are defined by

$$\psi_{\kappa\lambda}(p) := \frac{(2\lambda)^{\kappa+1/2}}{\sqrt{(2\kappa)!}} p^{\kappa+1/2} e^{-\lambda p}, \tag{159}$$

where κ is a positive integer and λ is a positive number with dimension of length. Using the formula

$$\int_0^\infty dp\, p^n e^{-\nu p} = \frac{n!}{\nu^{n+1}}, \tag{160}$$

which is valid for all non-negative integers n and for all complex ν that have a positive real part, we easily show that the wave packets are normalized,

$$\int_0^\infty \frac{dp}{p}\, \psi^2_{\kappa\lambda}(p) = 1. \tag{161}$$

Observe that $\psi_{\kappa\lambda}(p) = \psi_{\kappa 1}(\lambda p)$. The theory has been constructed so that it is scale invariant. The rescaling $p \mapsto \lambda p$ does not change the scalar product. The meaning of the parameter λ is, therefore, a scale: all energies concerning the packet will be proportional to λ^{-1}, all times and all radii to λ.

The expected energy,

$$\langle E\rangle_{\kappa\lambda} := \int_0^\infty \frac{dp}{p}\, p\psi^2_{\kappa\lambda}(p), \tag{162}$$

of the packet can be calculated by the formula (160) with the simple result

$$\langle E\rangle_{\kappa\lambda} = \frac{\kappa + 1/2}{\lambda}. \tag{163}$$

The (energy) width of the packet can be represented by the mean quadratic deviation (or dispersion),

$$\langle \Delta E\rangle_{\kappa\lambda} := \sqrt{\langle E^2\rangle_{\kappa\lambda} - \langle E\rangle^2_{\kappa\lambda}}\,, \tag{164}$$

which is

$$\langle \Delta E\rangle_{\kappa\lambda} = \frac{\sqrt{2\kappa+1}}{2\lambda}. \tag{165}$$

In the Schrödinger picture, the time evolution of the packet is generated by $-\hat{p}_t$:

$$\psi_{\kappa\lambda}(t,p) = \psi_{\kappa\lambda}(p) e^{-ipt}. \tag{166}$$

Let us calculate the corresponding wave function $\Psi_{\kappa\lambda}(r,t)$ in the r-representation,

$$\Psi_{\kappa\lambda}(t,r) := \int_0^\infty \frac{dp}{p}\ \psi_{\kappa\lambda}(t,p)\psi(r,p), \tag{167}$$

where the functions $\psi(r,p)$ are defined by (152). Formula (160) then yields:

$$\Psi_{\kappa\lambda}(t,r) = \frac{1}{\sqrt{2\pi}}\frac{\kappa!(2\lambda)^{\kappa+1/2}}{\sqrt{(2\kappa)!}}\left[\frac{i}{(\lambda+it+ir)^{\kappa+1}} - \frac{i}{(\lambda+it-ir)^{\kappa+1}}\right]. \tag{168}$$

It follows immediately that

$$\lim_{r\to 0}|\Psi_{\kappa\lambda}(t,r)|^2 = 0. \tag{169}$$

The scalar product measure for the r-representation is just dr because the eigenfunctions (152) are normalized, so the probability to find the shell between r and $r+dr$ is $|\Psi_{\kappa\lambda}(t,r)|^2 dr$.

Our first important result is, therefore, that the wave packets start away from the center $r=0$ and then are keeping away from it during the whole evolution. This can be interpreted as the *absence of singularity* in the quantum theory: no part of the packet is squeezed up to a point, unlike the shell in the classical theory.

Observe that the equation $\Psi_{\kappa\lambda}(t,0) = 0$ is *not* a result of an additional boundary condition imposed on the wave function. It follows from the dynamics we have constructed. The nature of the question that we are studying requires that the wave packets start in the asymptotic region so that their wave function vanishes at $r=0$ for $t\to-\infty$; this is the only condition put in by hand. The fact that the dynamics preserves this equation is the property of the self-adjoint extensions of the Hamiltonian and the position operators.

Again, λ can be used to re-scale the radius, $r=\lambda\rho$ and the time, $t=\lambda\tau$ with the result

$$\Psi_{\kappa\lambda}(\lambda\tau,\lambda\rho) = \frac{1}{\sqrt{\lambda}}\Psi_{\kappa 1}(\tau,\rho)\ . \tag{170}$$

The factor $1/\sqrt{\lambda}$ is due to the scalar product in the r-representation being

$$\big(\Phi(r),\Psi(r)\big) = \int_0^\infty dr\Phi^*(r)\Psi(r) = \int_0^\infty d\rho\sqrt{\lambda}\Phi^*(\rho)\sqrt{\lambda}\Psi(\rho)\ , \tag{171}$$

so that $\sqrt{\lambda}\Psi(\lambda\rho)$ is λ-independent.

A more tedious calculation is needed to obtain the time dependence $\langle r_t\rangle_{\kappa\lambda}$ of the expected radius of the shell,

$$\langle r_t\rangle_{\kappa\lambda} := \int_0^\infty dr\, r|\Psi_{\kappa\lambda}(t,r)|^2. \tag{172}$$

The results that can be calculated analytically are (we skip the calculations; for details, see [34]):

$$\langle r_t\rangle_{\kappa\lambda} = \langle r_{-t}\rangle_{\kappa\lambda} \quad \forall\kappa,\ \lambda\ ,t\ , \tag{173}$$

that is all packets are time reversal symmetric. There is the minimal expected radius $\langle r_0 \rangle_{\kappa\lambda}$ at $t = 0$,

$$\langle r_0 \rangle_{\kappa\lambda} = \frac{1}{\pi} \frac{2^{2\kappa}(\kappa!)^2}{(2\kappa)!} \frac{\kappa+1}{\kappa} \frac{\lambda}{\kappa+1/2} > 0 \ . \tag{174}$$

Let us turn to the asymptotics $t \to \pm\infty$. We obtain for both cases $t \to \pm\infty$ and all λ and κ:

$$\langle r_t \rangle_{\kappa\lambda} \approx |t| + O(t^{-2\kappa}). \tag{175}$$

We can also calculate the spread of the wave packet in r by means of the $\hat{r}$-dispersion $\langle \Delta r_t \rangle_{\kappa\lambda}$. The calculation of $\langle r^2 \rangle_{\kappa\lambda}$ is much easier than that of $\langle \hat{r} \rangle_{\kappa\lambda}$ because $\hat{r}^2$ is a differential operator in p-representation:

$$\langle r^2 \rangle_{\kappa\lambda} = -\lambda^2 \int_0^\infty \frac{dq}{q} \psi^*_{\kappa 1}(q,\tau) \left(-\sqrt{q} \frac{\partial^2}{\partial q^2} \frac{1}{\sqrt{q}} \psi_{\kappa 1}(q,\tau) \right) \ , \tag{176}$$

where $q = \lambda p$. This gives

$$\langle r_t^2 \rangle_{\kappa\lambda} = t^2 + \frac{\lambda^2}{2\kappa+1} \ . \tag{177}$$

For $\kappa \gg 1$, we can use the asymptotic expansion for the Γ-function to obtain

$$\frac{2^{2\kappa+1}(\kappa!)^2}{(2\kappa)!} = \sqrt{2\pi(2\kappa+1)} - \frac{1}{4}\sqrt{\frac{2\pi}{2\kappa+1}} + O(\kappa^{-3/2}) \ ; \tag{178}$$

that gives

$$\langle r_0 \rangle_{\kappa\lambda} = \lambda \left[\frac{1}{\sqrt{\pi\kappa}} + O(\kappa^{-3/2}) \right] \ . \tag{179}$$

Then, the spread of the packet for large κ is

$$\langle \Delta r_0 \rangle_{\kappa\lambda} = \lambda \left[\sqrt{\frac{\pi-2}{\kappa}} + O(\kappa^{-3/2}) \right] \ . \tag{180}$$

We observe that the spread is of the same order as the expected value. At $t \to \pm\infty$,

$$\langle \Delta r_{\pm\infty} \rangle_{\kappa\lambda} = \lambda \left[\frac{1}{\sqrt{2\kappa+1}} + O(t^{-2\kappa+1}) \right] \ . \tag{181}$$

Hence, the asymptotic spread is nearly equal to the spread at the minimum. This can be due to the light-like nature of the shell matter.

A further interesting question about the motion of the packets is about the portion of a given packet that moves in – is purely in-going – at a given time t. The portion is given by $\|\hat{P}_- \psi_{\kappa\lambda}\|^2$, where $\hat{P}_-$ is the projector defined in Sect. 4.1.

If we write out the projector kernel and make some simple rearrangements in the expression of the norm, we obtain:

$$\|\hat{P}_-\psi_{\kappa\lambda}\|^2 = \int_{-\infty}^{\infty} dq' \int_{-\infty}^{\infty} dq'' \left(\int_0^{\infty} da\, \psi_a^*(e^{q'})\psi_a(e^{q''})\right) \psi_{\kappa\lambda}^*(t, e^{q''})\psi_{\kappa\lambda}(t, e^{q'}), \tag{182}$$

where the transformation of integration variables p' and p'' to $e^{q'}$ and $e^{q''}$ in the projector kernels has been performed. The further calculation can be found in [34]. The results are:

$$\|\hat{P}_-\psi_{\kappa\lambda}\|^2_{t=0} = 1/2\ , \quad \|\hat{P}_-\psi_{\kappa\lambda}\|^2_{t\to-\infty} = 1\ , \quad \|\hat{P}_-\psi_{\kappa\lambda}\|^2_{t\to\infty} = 0\ . \tag{183}$$

Hence, we have one-to-one relation between in- and out-going states at the time of the bounce, while there are only in-going, or only out-going states at the infinity.

The obvious interpretation of these formulae is that the quantum shell always bounces at the center and re-expands.

The result that the quantum shell bounces and re-expands is clearly at odds with the classical idea of black hole forming in the collapse and preventing anything that falls into it from re-emerging. It is, therefore, natural to ask if the packet is squeezed enough so that an important part of it comes under its Schwarzschild radius. We can try to answer this question by comparing the minimal expected radius $\langle r_0\rangle_{\kappa\lambda}$ and its spread $\langle \varDelta r_0\rangle_{\kappa\lambda}$ with the expected Schwarzschild radius $\langle r_H\rangle_{\kappa\lambda}$ and its spread $\langle \varDelta r_H\rangle_{\kappa\lambda}$ for the wave packet. For the Schwarzschild radius, we have

$$\langle r_H\rangle_{\kappa\lambda} = 2\mathrm{G}\bar{M}_{\kappa\lambda} = (2\kappa+1)\frac{L_\mathrm{P}^2}{\lambda}\ , \tag{184}$$

where L_P is the Planck length, and

$$\langle \varDelta r_H\rangle_{\kappa\lambda} = 2\mathrm{G}\langle \varDelta M\rangle_{\kappa\lambda} = \sqrt{2\kappa+1}\frac{L_\mathrm{P}^2}{\lambda}\ . \tag{185}$$

We are to ask the question: for which λ and κ does the following inequality hold:

$$\langle r_0\rangle_{\kappa\lambda} + \langle \varDelta r_0\rangle_{\kappa\lambda} < \langle r_H\rangle_{\kappa\lambda} - \langle \varDelta r_H\rangle_{\kappa\lambda}\ . \tag{186}$$

If it holds, then the most of the packet becomes squeezed beyond the most of the Schwarzschild radius. From (179) and (180), we obtain

$$\frac{\lambda^2}{L_\mathrm{P}^2}\left(\frac{1}{\sqrt{\pi}} + \sqrt{\pi-2}\right) < (2\kappa+1)\sqrt{\kappa} - \sqrt{\kappa(2\kappa+1)}\ . \tag{187}$$

Clearly, if $\lambda \approx L_\mathrm{P}$, then the inequality holds for any large κ (starting from $\kappa = 2$). For larger λ, κ must be $(\lambda/L_\mathrm{P})^{4/3}\times$ larger. Hence, we must go to Planck regime in order that the packets cross their Schwarzschild radius.

To summarize: The packet can, in principle, fall under its Schwarzschild radius. Even in such a case, the packet bounces and re-expands.

This apparent paradox will be explained in the next section.

4.3 Grey Horizons

In this section, we try to explain the apparently contradictory result that the quantum shell can cross its Schwarzschild radius in both directions. The first possible idea that comes to mind is simply to disregard everything that our model says about Planck regime. This may be justified, because the model can hardly be considered as adequate for this regime. Outside Planck regime, however, the shell bounces before it reaches its Schwarzschild radius and there is no paradox. However, the model *is* mathematically consistent, simple, and solvable; it must, therefore, provide some mechanism to make the horizon leaky. We shall study this mechanism in the hope that it can work in more realistic situations, too.

To begin with, we have to recall that the Schwarzschild radius is the radius of a non-diverging null hyper-surface; anything moving to the future can cross such a hyper-surface only in one direction. The local geometry is that of an apparent horizon. (Whether or not an *event* horizon forms is another question; the answer to it also depends on the geometry near the singularity [42]). However, as Einstein's equations are invariant under time reversal, there are two types of Schwarzschild radius: that associated with a black hole and that associated with a white hole. Let us call these Schwarzschild radii (or apparent horizons) themselves *black* and *white*. The explanation of the paradox that follows from our model is that quantum states can contain a linear combination of black and white (apparent) horizons, and that no event horizon ever forms. We call such a combination a *grey horizon*.

The existence of grey horizons can be shown as follows. The position and the "colour" of a Schwarzschild radius outside the shell is determined by the spacetime metric. For our model, this metric is a combination of purely gauge and purely dependent degrees of freedom, and so it is determined, within the classical version of the theory, by the physical degrees of freedom through the constraints.

The metric can be calculated along a Cauchy surface Σ that intersect the shell with total energy E, direction of radial motion η, and the radius r. The result of this calculation (see [34]) is: If the shell is contracting, $\eta = -1$, then any space-like surface containing such a shell can at most intersect an *out-going* apparent horizon at the radius $R = 2\mathrm{G}E$. Analogous result holds for $\eta = +1$, where the shell is expanding: the apparent horizon is then in-going. The horizon radius is determined by the equation $R_H = 2\mathrm{G}E$, and the horizon will cut Σ if and only if $R_H > r$. We can assign the value $+1$ (-1) to the horizon that is out- (in-)going and denote the quantity by c (colour: black or white hole). To summarize:

1. The condition that an apparent horizon intersects Σ is $r < 2\mathrm{G}E$.
2. The position of the horizon at Σ is $R_H = 2\mathrm{G}E$.
3. The colour c of this horizon is $c = \eta$.

In this way, questions about the existence and colour of an apparent horizon outside the shell are reduced to equations containing dynamical variables of the shell. In particular, the result that $c = \eta$ can be expressed by saying that the shell always creates a horizon outside that cannot block its motion. All that matters is that the shell can bounce at the singularity (which it cannot within the classical theory).

These results can be carried over to quantum mechanics after quantities such as $2\mathrm{G}E - r$ and η are expressed in terms of the operators describing the shell. Then we obtain a "quantum horizon" with the "expected radius" $2\mathrm{G}\langle E\rangle$ and with the "expected colour" $\langle\eta\rangle$ to be mostly black at the time when the expected radius of the shell crosses the horizons inwards, neutrally grey at the time of the bounce and mostly white when the shell crosses it outwards.

This proof has, however, two weak points. First, the spacetime metric on the background manifold is not a gauge invariant quantity; although all gauge invariant geometrical properties can be extracted from it within the classical version of the theory, this does not seem to be possible in the quantum theory [43]. Second, calculating the quantum spacetime geometry along hyper-surfaces of a foliation on a given background manifold is foliation dependent. For example, one can easily imagine two hyper-surfaces Σ and Σ' belonging to different foliations, that intersect each other at a sphere outside the shell and such that Σ intersects the shell in its in-going and Σ' in its out-going state. Observe that the need for a foliation is only due to our insistence on calculating the quantum metric on the background manifold.

The essence of these problems is the gauge dependence of the results of the calculation. However, it seems that this dependence concerns only details such as the distribution of different hues of grey along the horizon, not the qualitative fact that the horizon exists and changes colour from almost black to almost white. Still, a more reliable method to establish the existence and properties of grey horizons might require another material system to be coupled to our model; this could probe the spacetime geometry around the shell in a gauge-invariant way.

4.4 Concluding Remarks

Comparison of the motion of wave packets of Sect. 4.2 with the classical dynamics of the shell as described in Sect. 3.1 shows a notable difference. Whereas all classical shells cross their Schwarzschild radius and reach the singularity in some stage of their evolution, the quantum wave packets never reach the singularity, but always bounce and re-expand; some of them even manage to cross their Schwarzschild radius during their motion. This behaviour is far from being a small perturbation around a classical solution if the classical spacetime is considered as a whole. Even locally, the semi-classical approximation is not valid near the bouncing point. It is surely valid in the whole asymptotic region, where narrow wave packets follow more or less the classical trajectories of the shell.

The most important question, however, concerns the validity of the semi-classical approximation near the Schwarzschild radius. We have seen that the geometry near the radius can resemble the classical black hole geometry in the neighbourhood of the point where the shell is crossing the Schwarzschild radius inwards. Then, the radius changes its colour gradually and the geometry becomes very different from the classical one. Finally, near the point where the shell crosses the Schwarzschild radius outwards, the radius is predominantly white and the quantum geometry can be again similar to the classical geometry, this time of a white hole horizon.

If the change of colour is very slow, then the neighbourhood of the inward crossing where the classical geometry is a good approximation can be large. It seems that a sufficiently large scattering time would allow for arbitrarily slow change of colour. We cannot exclude, therefore, that the quantum spacetime contains an extended region with the geometry resembling its classical counterpart near a black hole horizon. This can be true even if the quantum spacetime as a whole differs strongly from any typical classical collapse solution.

One can even imagine the following scenario (which needs a more realistic model than a single thin shell). A quantum system with a large energy collapses and re-expands after a huge scattering time. The black hole horizon phase is so long that Hawking evaporation becomes significant and must be taken into account in the calculation. It does then influence the scattering time and the period of validity of the black hole approximation. The black hole becomes very small and only then the change of horizon colour becomes significant. The white hole stage is quite short and it is only the small remnant of the system that, finally, re-expands. The whole process can still preserve unitarity. In fact, this is a scenario for the issue of Hawking evaporation process. It is not excluded by the results of the present paper.

One can also consider the following conception. At and under the Schwarzschild radius, the local spacetime geometry for a white Schwarzschild radius is very, and measurably, different from that of a black one. However, outside the Schwarzschild radius, the local geometries of both cases are isometric to each other so that the isometry need not contain the time reversal. Hence, the quantum geometry outside a very grey horizon need not actually differ from the classical geometry around a black hole (locally) very much.

All these speculations must be investigated and made more precise. If an observer staying at the radius R_B measures the proper time $\Delta T(R_B)$ between the shell wave packet crossing his position in and out, then of course $\lim_{R_B \to \infty} \Delta T(R_B) = \infty$. In the collision theory, one considers, therefore, some standard collision with $\Delta_s T(R_B)$ and takes the limit only after the subtraction,

$$\lim_{R_B \to \infty} \left(\Delta T(R_B) - \Delta_s T(R_B) \right) . \tag{188}$$

This limit, if finite, is called *time delay* [44]. A particularly difficult example is Coulomb scattering, for which the divergence in $\Delta T(R_B)$ involves not only linear, but also logarithmic terms. This is due to the long range of Coulomb potential and is, therefore, called *infrared divergence*. One can still find a subtraction for the Coulomb case because the logarithmic term depends only on charges and so has the same leading part for a whole class of scattering processes.

In our case, there are also logarithmic terms, but they depend on the energy of the shell (energy is the charge for gravitation). No reasonable subtraction seems, therefore, possible. One had better work at a finite radius R_B all time.

To calculate our $\Delta T(R_B)$, one has to use the quantum geometry outside the shell.

But what is the quantum geometry? In the classical version of general relativity, a metric tensor field written in particular coordinates contains all information about local geometric properties. A change in the coordinates does not lead to any change of these properties because they can be calculated from the transformed metric. In a quantum theory, however, coordinate transformations is a delicate issue. Let us turn to our simple system to see why. We have chosen the coordinates U and V that are uniquelly determined for each solution by means of their geometric properties. The components $A(U,V)$ and $R(U,V)$ of the four-metric with respect to these coordinates are geometric quantities themselves. Nobody wants to deny that the quantities $A(\eta,M,w;U,V)$ and $R(\eta,M,w;U,V)$ are Dirac observables for each fixed value of U and V: they are then just functions of η, M and w, which *are* Dirac observables. In the quantum theory, one can try to promote these quantities to operators by replacing η, M and w by the corresponding quantum operators and by choosing a suitable factor ordering. Let us suppose that the quantities $\hat{A}(U,V)$ and $\hat{R}(U,V)$ obtained in this way are well defined operators for each value of U and V. Can we not calculate all geometric properties near each point (U,V) from these operators?

Problems emerge if we choose to work with a different set of geometric coordinates. Let us, for example, pass to the Schwarzschild coordinates T and R; near infinity, they are a very natural choice. Again, we can work out the form of the metric components $g_{00}(\eta,M,w;T,R)$ and $g_{11}(\eta,M,w;T,R)$ for each solution and try to construct the operators $\hat{g}_{00}(T,R)$ and $\hat{g}_{11}(T,R)$ from them. Is there any transformation within the quantum theory analogous to the classical coordinate transformation and providing a basis for a proof that the local geometric properties calculated from $\hat{A}(U,V)$ and $\hat{R}(U,V)$ are the same as those calculated from $\hat{g}_{00}(T,R)$ and $\hat{g}_{11}(T,R)$? A problem is that the transformation between $\{U,V\}$ and $\{T,R\}$ is field dependent. Indeed, for $\eta = +1$, it is given by (70) and (71). The right-hand sides depend not only on T_∞ (which coincides with our T here) and R, but also on M and w (the full transformation depends also on η). Hence, if we assume that the quantities T and R are just parameters, then the quantities U and V are operators and vice versa. However, the quantities T and R are parameters in the operators $\hat{g}_{00}(T,R)$ and $\hat{g}_{11}(T,R)$ and the quantities U and V are parameters in $\hat{A}(U,V)$ and $\hat{R}(U,V)$.

It seems, therefore, that there is a problem of principle in addition to the factor-ordering problem to transform the operators $\hat{A}(U,V)$ and $\hat{R}(U,V)$ to $\hat{g}_{00}(T,R)$ and $\hat{g}_{11}(T,R)$. The conclusion is that quantum geometry cannot be described in a coherent manner analogous to the differential geometry, see also [43].

At the present stage, we try to calculate each geometric property for itself. Thus, in his PhD thesis, M. Ambrus, to whom I owe all my knowledge about collision theory, tries to define and calculate the scattering times. Another attempt is to define quantum geometry in a similar way as the classical geometry is defined: by properties of test particles. This might work even near the

Schwarzschild and zero radius. Thus, we are trying with I. Kouletsis to quantize an analogous system containing *two* shells and to use the second shell as a spy probing the quantum geometry created by the first one. For some preliminary results see [45], [28] and [46].

The calculations of this paper are valid only for null shells. Similar calculations for massive shells have been performed in [47]. There has been re-expansion and unitarity for massive shells if the rest mass has been smaller than the Planck mass (10^{-5} g). It is very plausible that the interpretation of these results is similar to that given in the present paper. Thus, we can expect the results valid at least for all "light" shells. There is, in any case, a long way to any astrophysically significant system and a lot of work is to be done before we can claim some understanding of the collapse problem.

Acknowledgments

Discussions with M. Ambrus, A. Ashtekar, A. Barvinski, R. Beig, H. Friedrich, T. Jacobson, J. Jezierski, C. Kiefer, J. Kijowski, I. Kouletsis, K. Kuchař, J. Louko and C. Vaz have strongly helped me to form the ideas of these lectures.

References

1. P.A.M. Dirac: *Proc. Roy. Soc. London A* **268**, 57 (1962)
2. V.P. Frolov and G.A. Vilkovisky: *Phys. Lett. B* **106**, 307 (1981)
3. P. Hájíček: *J. Math. Phys.* **36**, 4612 (1995)
4. P. Hájíček: *Class. Quant. Grav.* **13**, 1353 (1996)
5. P. Hájíček and J. Kijowski: *Phys. Rev. D* **61**, 024037 (2000)
6. P. Hájíček and C. Kiefer: *Nucl. Phys. B* **603**, 531 (2001)
7. S.W. Hawking and G.F.R. Ellis: *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, 1973)
8. P.T. Chrusciel, J. Jezierski and J. Kijowski: *Hamiltonian Field Theory in the Radiating Regime* (Springer-Verlag, Berlin, 2002)
9. R. Beig and N. O'Murchadha: *Ann. Phys.* (N.Y.) **174**, 463 (1987)
10. R. Arnowitt, S. Deser and C. W. Misner: The dynamics of general relativity, in L. Witten (ed.), *Gravitation: An Introduction to Current research* (Wiley, New York, 1962)
11. A.E. Fischer and J.E. Marsden: in S.W. Hawking and W. Israel (eds.): *General Relativity. An Einstein Centenary Review* (Cambridge University Press, Cambridge, 1979)
12. C. Teitelboim: *Ann. Phys.* (N.Y.) **79**, 542 (1973)
13. I. Kouletsis: Thesis, `gr-qc/9801019`
14. P.A.M. Dirac: *Lectures on Quantum Mechanics* (Yeshiva University Press, New York, 1964)
15. M. Henneaux and C. Teitelboim: *Quantization of Gauge Systems* (Princeton University Press, Princeton N.J., 1992)
16. T. Regge and C. Teitelboim: *Ann. Phys.* (N.Y.) **88**, 286 (1974)
17. K. Grabowska and J. Kijowski: *Canonical Gravity and Gravitational Energy*, preprint, Warsaw, 2002

18. P. Bergmann: *Rev. Mod. Phys.* **33**, 510 (1961)
19. K.V. Kuchař: in: R.J. Gleiser, C.N. Kozameh and O.M. Moreschi (eds.), *General Relativity and Gravitation 1992* (Institute of Physics, Bristol, 1992)
20. C.G. Torre: *Phys. Rev. D* **48** (1993) R2373
21. B.C. DeWitt: *Dynamical Theory of Groups and Fields* (Gordon and Breach, New York, 1965)
22. A. Ashtekar: *Asymptotic Quantization* (Bibliopolis, Napoli, 1987)
23. K.V. Kuchař: *Phys. Rev. D* **4**, 955 (1971)
24. K.V. Kuchař: *J. Math. Phys.* **13**, 768 (1972)
25. P.A.M. Dirac: *Rev. Mod. Phys.* **21**, 392 (1949)
26. J. Kijowski and W.M. Tulczyjew: *A Symplectic Framework for Field Theories*, Lecture Notes in Physics, Vol. 107 (Springer-Verlag, Berlin 1979).
27. C. Rovelli: *Phys. Rev. D* **42** (1990) 2638; D **43** (1991) 442; D **44** (1991) 1339
28. P. Hájíček and I. Kouletsis: *Class. Quantum Grav.* **19**, 2551 (2002)
29. C. Barrabès and W. Israel: *Phys. Rev. D* **43**, 1129 (1991)
30. J. Louko, B. Whiting, and J. Friedman: *Phys. Rev. D* **57**, 2279 (1998)
31. P. Kraus and F. Wilczek: *Nucl. Phys. B* **433**, 403 (1995)
32. K.V. Kuchař: *Phys. Rev. D* **50**, 3961 (1994)
33. C.G. Torre and I.M. Anderson: *Phys. Rev. Lett.* **70**, 3525 (1993)
34. P. Hájíček: *Nucl. Phys. B* **603**, 555 (2001)
35. C.J. Isham: in: B.S. DeWitt and R. Stora (eds.), *Relativity, Groups and Topology II* (Elsevier, Amsterdam, 1984)
36. C. Rovelli: *Nuovo Cim. B* **100**, 343 (1987)
37. P. Hájíček: in: J. Ehlers and H. Friedrich (eds.), *Canonical Gravity: From Classical to Quantum* (Springer-Verlag, Berlin, 1994)
38. P. Hájíček and C.J. Isham: *J. Math. Phys.* **37**, 3522 (1996); P. Hájíček: *J. Math. Phys.* **39**, 4824 (1998).
39. A.O. Barut and R. Rączka: *Theory of Group Representations and Applications* (Polish Scientific Publishers, Warsaw, 1980)
40. T.D. Newton and E.P. Wigner: *Rev. Mod. Phys.* **21**, 400 (1949)
41. M. Reed, and B. Simon: *Methods of Modern Mathematical Physics, Vol. 2. Fourier Analysis, Self-Adjointness* (Academic Press, New York, 1975)
42. P. Hájíček: *Phys. Rev. D* **36**, 1065 (1987)
43. P. Hájíček: *Nucl. Phys. B* (Proc. Suppl.) **80** (2000), `gr-qc/9903089`
44. J.D. Dollard: *J. Math. Phys.* **5**, 729 (1964)
45. P. Hájíček and I. Kouletsis: *Class. Quantum Grav.* **19**, 2529 (2002)
46. I. Kouletsis and P. Hájíček: *Class. Quantum Grav.* **19**, 2567 (2002)
47. P. Hájíček, B. S. Kay, and K.V. Kuchař: *Phys. Rev. D* **46**, 5439 (1992)

Primordial Black Holes as a Probe of Cosmology and High Energy Physics

Bernard J. Carr

Astronomy Unit, Queen Mary, University of London,
Mile End Road, London E1 4NS, England

Abstract. Recent developments in the study of primordial black holes (PBHs) will be reviewed, with particular emphasis on their formation and evaporation. PBHs could provide a unique probe of the early Universe, gravitational collapse, high energy physics and quantum gravity. Indeed their study may place interesting constraints on the physics relevant to these areas even if they never formed.

1 Introduction

Hawking's discovery in 1974 that black holes emit thermal radiation due to quantum effects was surely one of the most important results in 20th century physics. This is because it unified three previously disparate areas of physics - quantum theory, general relativity and thermodynamics - and like all such unifying ideas it has led to profound insights. Although not strictly an application of quantum gravity theory, the theme of this meeting, it might be regarded as a conceptual first step in that direction. Also there is a natural link in that the final stage of black hole evaporation, when the black hole is close to the Planck mass, can only be understood with a proper theory of quantum gravity.

In practice, only "primordial black holes" which formed in the early Universe could be small enough for Hawking radiation to be important. Such a black hole will be referred to by the acronym "PBH", although this should not be confused with the acronym for "Physikzentrum Bad Honnef", the institute hosting this meeting! Interest in PBHs goes back nearly 35 years and some of the history of the subject will be reviewed in Sect. 2. As will be seen, interest was much intensified as a result of Hawking's discovery. Indeed, although it is still not definite that PBHs ever formed, it was only through thinking about them that Hawking was led to his remarkable insight. Thus the discovery illustrates that studying something may be useful even if it does not exist!

Of course, the subject is much more interesting if PBHs *did* form and their discovery would provide a unique probe of at least four areas of physics: the early Universe; gravitational collapse; high energy physics; and quantum gravity. The first topic is relevant because studying PBH formation and evaporation can impose important constraints on primordial inhomogeneities, cosmological phase transitions (including inflation) and varying–G models. These topics are covered in Sects. 3, 4, and 5, respectively. The second topic is discussed in Sect. 6 and relates to recent developments in the study of "critical phenomena" and the issue of whether PBHs are viable dark matter candidates. The third topic arises

B.J. Carr, Primordial Black Holes as a Probe of Cosmology and High Energy Physics, Lect. Notes Phys. **631**, 301–321 (2003)
http://www.springerlink.com/

because PBH evaporations could contribute to cosmic rays, whose energy distribution would then give significant information about the high energy physics involved in the final explosive phase of black hole evaporation. This is covered in Sect. 7. The fourth topic arises because it has been suggested that quantum gravity effects could appear at TeV scale and this leads to the intriguing possibility that small black holes could be generated in accelerators experiments or cosmic ray events. As discussed in Sect. 8, this could have striking observational consequences. Although such black holes are not technically "primordial", this possibility would have radical implications for PBHs themselves.

2 Historical Overview

It was realized many years ago that black holes with a wide range of masses could have formed in the early Universe as a result of the great compression associated with the Big Bang. A comparison of the cosmological density at a time t after the Big Bang with the density associated with a black hole of mass M shows that PBHs would have of order the particle horizon mass at their formation epoch:

$$M_H(t) \approx \frac{c^3 t}{G} \approx 10^{15} \left(\frac{t}{10^{-23}\ \mathrm{s}} \right) \mathrm{g}. \tag{1}$$

PBHs could thus span an enormous mass range: those formed at the Planck time (10^{-43}s) would have the Planck mass (10^{-5}g), whereas those formed at 1 s would be as large as $10^5 M_\odot$, comparable to the mass of the holes thought to reside in galactic nuclei. By contrast, black holes forming at the present epoch could never be smaller than about $1 M_\odot$.

Zeldovich & Novikov [119] first derived (1) but they were really considering "retarded cores" rather than black holes and Hawking [54] was the first person to realize that primordial density perturbations might lead to gravitational collapse on scales above the Planck mass. For a while the existence of PBHs seemed unlikely since Zeldovich & Novikov [119] had pointed out that they might be expected to grow catastrophically. This is because a simple Newtonian argument suggests that, in a radiation-dominated universe, black holes much smaller than the horizon cannot grow much at all, whereas those of size comparable to the horizon could continue to grow at the same rate as it throughout the radiation era. Since we have seen that a PBH must be of order the horizon size at formation, this suggests that all PBHs could grow to have a mass of order $10^{15} M_\odot$ (the horizon mass at the end of the radiation era). There are strong observational limits on how many such giant holes the Universe could contain, so the implication seemed to be that very few PBHs ever existed.

However, the Zeldovich-Novikov argument was questionable since it neglected the cosmological expansion and this would presumably hinder the black hole growth. Indeed myself and Hawking were able to disprove the notion that PBHs could grow at the same rate as the particle horizon by demonstrating that there is no spherically symmetric similarity solution which represents a black hole attached to an exact Friedmann model via a sound-wave [22]. Since a PBH must therefore soon become much smaller than the horizon, at which stage cosmological effects become unimportant, we concluded that PBHs cannot grow very much at all (cf. [12,80]).

The realization that small PBHs might exist after all prompted Hawking to study their quantum properties. This led to his famous discovery [55] that black holes radiate thermally with a temperature

$$T = \frac{\hbar c^3}{8\pi GMk} \approx 10^{-7} \left(\frac{M}{M_\odot}\right)^{-1} \text{ K}, \tag{2}$$

so they evaporate on a timescale

$$\tau(M) \approx \frac{G^2 M^3}{\hbar c^4} \approx 10^{64} \left(\frac{M}{M_\odot}\right)^{3} \text{ y}. \tag{3}$$

Only black holes smaller than 10^{15}g would have evaporated by the present epoch, so (1) implies that this effect could be important only for black holes which formed before 10^{-23}s.

Despite the conceptual importance of this result, it was bad news for PBH enthusiasts. For since PBHs with a mass of 10^{15}g would be producing photons with energy of order 100 MeV at the present epoch, the observational limit on the γ-ray background intensity at 100 MeV immediately implied that their density could not exceed 10^{-8} times the critical density [101]. Not only did this render PBHs unlikely dark matter candidates, it also implied that there was little chance of detecting black hole explosions at the present epoch [103]. Nevertheless, it was realized that PBH evaporations could still have interesting cosmological consequences. In particular, they might generate the microwave background [120] or modify the standard cosmological nucleosynthesis scenario [98] or contribute to the cosmic baryon asymmetry [3]. PBH evaporations might also account for the annihilation-line radiation coming from the Galactic centre [99] or the unexpectedly high fraction of antiprotons in cosmic rays [73]. PBH explosions occurring in an interstellar magnetic field might also generate radio bursts [105]. Even if PBHs had none of these consequences, studying such effects leads to strong upper limits on how many of them could ever have formed and thereby constrains models of the early Universe.

Originally it was assumed that PBHs would form from initial inhomogeneities but in the 1980s attention switched to several new formation mechanisms. Most of the mechanisms were associated with various phase transitions that might be expected to occur in the early Universe and there was particular interest in whether PBHs could form from the quantum fluctuations associated with the many different types of inflationary scenarios. Indeed it soon became clear that there are many ways in PBHs serve as a probe of the early Universe and, even if they never formed, their non-existence gives interesting information [20]. In this sense, they are similar to other “relicts” of the Big Bang, except that they derive from much earlier times.

In the 1990s work on the cosmological consequences of PBH evaporations was revitalized as a result of calculations by my PhD student Jane MacGibbon. She realized that the usual assumption that particles are emitted with a black-body spectrum as soon as the temperature of the hole exceeds their rest mass is too simplistic. If one adopts the conventional view that all particles are composed of a small number of fundamental point-like constituents (quarks and leptons), it would seem natural to assume that it is these fundamental particles rather than the composite ones which are emitted directly once the temperature goes above the QCD confinement scale of 250 MeV. One can therefore envisage a black hole as emitting relativistic quark and gluon jets which subsequently fragment into leptons and hadrons [83,85] and this modifies the cosmological constraints considerably [84]

Over the last decade PBHs have been assigned various other cosmological roles. Some people have speculated that PBH evaporation, rather than proceeding indefinitely, could cease when the black hole gets down to the Planck mass [13,30]. In this case, one could end up with stable Planck mass relics, which would provide dark matter candidates [7,25,82]. Although most gamma-ray bursts are now known to be at cosmological distances, it has been proposed that some of the short period ones could be nearby exploding PBHs [10,28]. Solar mass PBHs could form at the quark-hadron phase transition and, since some of these should today reside in our Galactic halo, these have been invoked to explain the microlensing of stars in the Magellanic Clouds [64,66,115].

3 PBHs as a Probe of Primordial Inhomogeneities

One of the most important reasons for studying PBHs is that it enables one to place limits on the spectrum of density fluctuations in the early Universe. This is because, if the PBHs form directly from density perturbations, the fraction of regions undergoing collapse at any epoch is determined by the root-mean-square amplitude ϵ of the fluctuations entering the horizon at that epoch and the equation of state $p = \gamma\rho$ $(0 < \gamma < 1)$. One usually expects a radiation equation of state $(\gamma = 1/3)$ in the early Universe. In order to collapse against the pressure, an overdense region must be larger than the Jeans length at maximum expansion and this is just $\sqrt{\gamma}$ times the horizon size. On the other hand, it cannot be larger than the horizon size, else it would form a separate closed universe and not be part of our Universe [22].

This has two important implications. Firstly, PBHs forming at time t should have of order the horizon mass given by (1). Secondly, for a region destined to collapse to a PBH, one requires the fractional overdensity at the horizon epoch δ to exceed γ. Providing the density fluctuations have a Gaussian distribution and are spherically symmetric, one can infer that the fraction of regions of mass M which collapse is [18]

$$\beta(M) \sim \epsilon(M) \exp\left[-\frac{\gamma^2}{2\epsilon(M)^2}\right] \tag{4}$$

where $\epsilon(M)$ is the value of ϵ when the horizon mass is M. The PBHs can have an extended mass spectrum only if the fluctuations are scale-invariant (i.e. with

ϵ independent of M). In this case, the PBH mass spectrum is given by [18]

$$dn/dM = (\alpha - 2)(M/M_*)^{-\alpha} M_*^{-2} \Omega_{\rm PBH} \rho_{\rm crit} \,, \tag{5}$$

where $M_* \approx 10^{15}$g is the current lower cut-off in the mass spectrum due to evaporations, $\Omega_{\rm PBH}$ is the total density of the PBHs in units of the critical density (which itself depends on β) and the exponent α is determined by the equation of state:

$$\alpha = \left(\frac{1+3\gamma}{1+\gamma}\right) + 1 \,. \tag{6}$$

$\alpha = 5/2$ if one has a radiation equation of state (γ=1/3), as expected. This means that the integrated mass density of PBHs larger than M falls off as $M^{-1/2}$, so most of the PBH density is contained in the smallest ones.

Many scenarios for the cosmological density fluctuations predict that ϵ is at least approximately scale-invariant but the sensitive dependence of β on ϵ means that even tiny deviations from scale-invariance can be important. If $\epsilon(M)$ decreases with increasing M, then the spectrum falls off exponentially and most of the PBH density is contained in the smallest ones. If $\epsilon(M)$ increases with increasing M, the spectrum rises exponentially and - if PBHs were to form at all - they could only do so at large scales. However, the microwave background anisotropies would then be larger than observed, so this possibilty can be rejected.

The current density parameter $\Omega_{\rm PBH}$ associated with PBHs which form at a redshift z or time t is related to β by [18]

$$\Omega_{\rm PBH} = \beta \Omega_{\rm R} (1+z) \approx 10^6 \beta \left(\frac{t}{\rm s}\right)^{-1/2} \approx 10^{18} \beta \left(\frac{M}{10^{15}{\rm g}}\right)^{-1/2} \,, \tag{7}$$

where $\Omega_{\rm R} \approx 10^{-4}$ is the density parameter of the microwave background and we have used (1). The $(1+z)$ factor arises because the radiation density scales as $(1+z)^4$, whereas the PBH density scales as $(1+z)^3$. Any limit on $\Omega_{\rm PBH}$ therefore places a constraint on $\beta(M)$ and the constraints are summarized in Fig. 1, which is taken from Carr et al. [25]. The constraint for non-evaporating mass ranges above 10^{15}g comes from requiring $\Omega_{\rm PBH} < 1$ but stronger constraints are associated with PBHs smaller than this since they would have evaporated by now [19]. The strongest one is the γ-ray limit associated with the 10^{15}g PBHs evaporating at the present epoch [101]. Other ones are associated with the generation of entropy and modifications to the cosmological production of light elements [98]. The constraints below 10^6g are based on the (uncertain) assumption that evaporating PBHs leave stable Planck mass relics, in which case these relics are required to have less than the critical density [7,25,82].

The constraints on $\beta(M)$ can be converted into constraints on $\epsilon(M)$ using (4) and these are shown in Fig. 2. Also shown here are the (non-PBH) constraints associated with the spectral distortions in the cosmic microwave background induced by the dissipation of intermediate scale density perturbations and the COBE quadrupole measurement. This shows that one needs the fluctuation amplitude to decrease with increasing scale in order to produce PBHs and the lines corresponding to various slopes in the $\epsilon(M)$ relationship are also shown in Fig. 2.

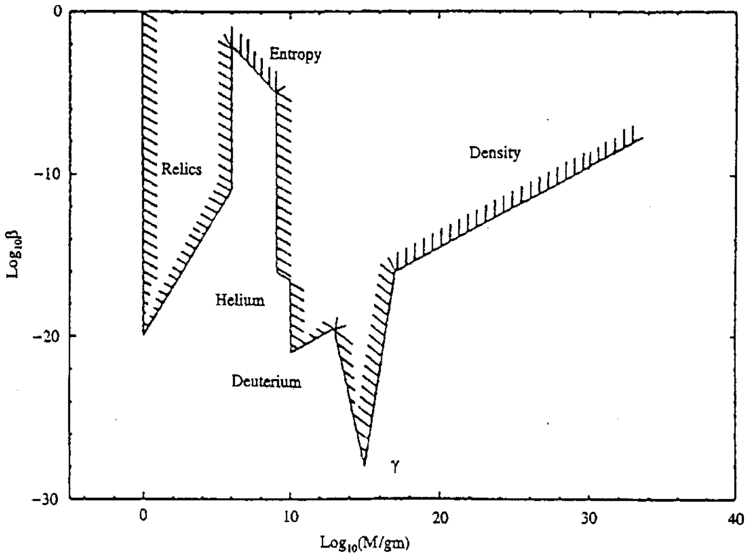


Fig. 1. Constraints on $\beta(M)$

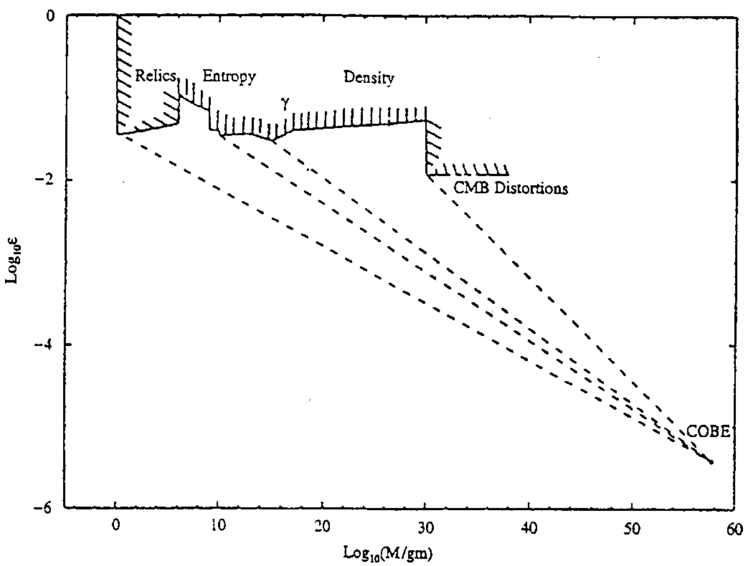


Fig. 2. Constraints on $\epsilon(M)$

4 PBHs as Probe of Cosmological Phase Transitions

Many phase transitions could occur in the early Universe which lead to PBH formation. Some of these mechanisms still require pre-existing density fluctuations but in others the PBHs form spontaneously even if the Universe starts off perfectly smooth. In the latter case, $\beta(M)$ depends not on $\epsilon(M)$ but on some other cosmological parameter.

4.1 Soft Equation of State

Some phase transitions can lead to the equation of state becoming soft ($\gamma << 1$) for a while. For example, the pressure may be reduced if the Universe's mass is ever channelled into particles which are massive enough to be non-relativistic. In such cases, the effect of pressure in stopping collapse is unimportant and the probability of PBH formation just depends upon the fraction of regions which are sufficiently spherical to undergo collapse; this can be shown to be [70]

$$\beta = 0.02\, \epsilon^{13/2}\,. \tag{8}$$

The value of β is now much less sensitive to ϵ than indicated by (4) and most of the PBHs will be smaller than the horizon mass at formation by a factor $\epsilon^{3/2}$. For a given spectrum of primordial fluctuations, this means that there may just be a narrow mass range - associated with the period of the soft equation of state - in which the PBHs form. In particular, this could happen at the quark-hadron phase transition since the pressure may then drop for a while [66].

4.2 Collapse of Cosmic Loops

In the cosmic string scenario, one expects some strings to self-intersect and form cosmic loops. A typical loop will be larger than its Schwarzschild radius by the inverse of the factor $G\mu$, where μ is the mass per unit length. If strings play a role in generating large-scale structure, $G\mu$ must be of order 10^{-6}. Hawking [57] showed that there is always a small probability that a cosmic loop will get into a configuration in which every dimension lies within its Schwarzschild radius and he estimated this to be

$$\beta \sim (G\mu)^{-1}(G\mu x)^{2x-2}\,, \tag{9}$$

where x is the ratio of the loop length to the correlation scale. If one takes x to be 3, $\Omega_{\rm PBH} > 1$ for $G\mu > 10^{-7}$, so he argued that one overproduces PBHs in the favoured string scenario. Polnarev & Zemboricz [102] obtained a similar result. However, $\Omega_{\rm PBH}$ is very sensitive to x and a slight reduction could still give an interesting value [17,41,86]. Note that spectrum (5) still applies since the holes are forming with equal probability at every epoch.

4.3 Bubble Collisions

Bubbles of broken symmetry might arise at any spontaneously broken symmetry epoch and various people, including Hawking, suggested that PBHs could form as a result of bubble collisions [32,58,78]. However, this happens only if the bubble formation rate per Hubble volume is finely tuned: if it is much larger than the Hubble rate, the entire Universe undergoes the phase transition immediately and there is not time to form black holes; if it is much less than the Hubble rate, the bubbles are very rare and never collide. The holes should have a mass of order the horizon mass at the phase transition, so PBHs forming at the GUT epoch would have a mass of 10^3g, those forming at the electroweak unification epoch would have a mass of 10^{28}g, and those forming at the QCD (quark-hadron) phase transition would have mass of around $1M_\odot$. Only a phase transition before 10^{-23}s would be relevant in the context of evaporating PBHs.

4.4 Inflation

Inflation has two important consequences for PBHs. On the one hand, any PBHs formed before the end of inflation will be diluted to a negligible density. Inflation thus imposes a lower limit on the PBH mass spectrum:

$$M > M_{\rm min} = M_{\rm Pl} \left(\frac{T_{\rm RH}}{T_{\rm Pl}} \right)^{-2} , \tag{10}$$

where $T_{\rm RH}$ is the reheat temperature and $T_{\rm Pl} \approx 10^{19}$ GeV is the Planck temperature. The CMB quadrupole measurement implies $T_{\rm RH} \approx 10^{16}$GeV, so $M_{\rm min}$ certainly exceeds 1g. On the other hand, inflation will itself generate fluctuations and these may suffice to produce PBHs after reheating. If the inflaton potential is $V(\phi)$, then the horizon-scale fluctuations for a mass-scale M are

$$\epsilon(M) \approx \left[\frac{V^{3/2}}{M_{\rm Pl}^3 V'} \right]_H \tag{11}$$

where a prime denotes $d/d\phi$ and the right-hand-side is evaluated for the value of ϕ when the mass-scale M falls within the horizon.

In the standard chaotic inflationary scenario, one makes the "slow-roll" and "friction-dominated" asumptions:

$$\xi \equiv \left(\frac{M_{\rm Pl} V'}{V} \right)^2 \ll 1, \quad \eta \equiv \frac{M_{\rm Pl}^2 V''}{V} \ll 1 \, . \tag{12}$$

Usually the exponent n characterizing the power spectrum of the fluctuations, $|\delta_k|^2 \approx k^n$, is very close to but slightly below 1:

$$n = 1 + 4\xi - 2\eta \approx 1. \tag{13}$$

Since ϵ scales as $M^{(1-n)/4}$, this means that the fluctuations are slightly increasing with scale. The normalization required to explain galaxy formation ($\epsilon \approx 10^{-5}$)

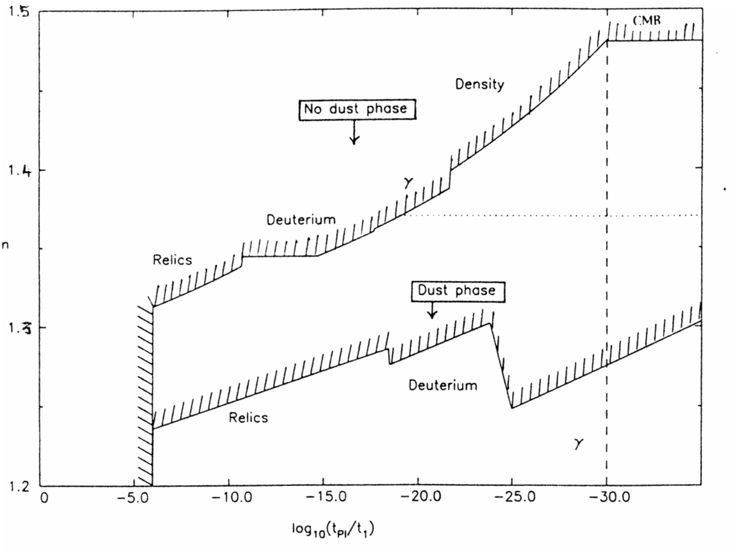


Fig. 3. Constraints on spectral index n in terms of reheat time t_1.

would then preclude the formation of PBHs on a smaller scale. If PBH formation is to occur, one needs the fluctuations to decrease with increasing mass ($n > 1$) and this is only possible if the scalar field is accelerating sufficiently fast:

$$V''/V > (1/2)(V'/V)^2. \tag{14}$$

This condition is certainly satisfied in some scenarios [23] and, if it is, (4) implies that the PBH density will be dominated by the ones forming immediately after reheating. Since each value of n corresponds to a straight line in Fig. 2, any particular value for the reheat time t_1 corresponds to an upper limit on n. This limit is indicated in Fig. 3, which is taken from Carr et al. [25] apart from a correction pointed out by Green & Liddle [47]. Similar constraints have now been obtained by several other people [15,72]. The figure also shows how the constraint on n is strengthened if the reheating at the end of inflation is sufficiently slow for there to be a dust-like phase [49]. PBHs have now been used to place constraints on many other sorts of inflationary scenarios - supernatural [104], supersymmetric [44], hybrid [40,68], oscillating [110], preheating [9,34,38,50] and running mass [79] - as well as a scenarios in which the inflaton serves as the dark matter [81].

Bullock & Primack [16] and Ivanov [63] have questioned whether the Gaussian assumption which underlies (4) is valid in the context of inflation. So long as the fluctuations are small ($\delta\phi/\phi << 1$), as certainly applies on a galactic scale, this assumption is valid. However, for PBH formation one requires $\delta\phi/\phi \sim 1$, and, in this case, the coupling of different Fourier modes destroys the Gaussianity. Their analysis suggests that $\beta(M)$ is much less than indicated by (4) but it still depends very sensitively on ϵ.

Not all inflationary scenarios predict that the spectral index should be constant. Hodges & Blumenthal [61] pointed out that one can get any form for the fluctuations whatsoever by suitably choosing the form of $V(\phi)$. For example, (11) suggests that one can get a spike in the spectrum by flattening the potential over some mass range (since the fluctuation diverges when V' goes to 0). This idea was exploited by Ivanov et al. [64], who fine-tuned the position of the spike so that it corresponds to the microlensing mass-scale.

5 PBHs as a Probe of a Varying Gravitational Constant

The PBH constraints would be severely modified if the value of the gravitational "constant" G was different at early times. The simplest varying-G model is Brans-Dicke (BD) theory [14], in which G is associated with a scalar field ϕ and the deviations from general relativity are specified by a parameter ω. A variety of astrophysical tests currently require $|\omega| > 500$, which implies that the deviations can only ever be small [113]. However, there exist generalized scalar-tensor theories [11,97,112] in which ω is itself a function of ϕ and these lead to a considerably broader range of variations in G. In particular, it permits ω to be small at early times (allowing noticeable variations of G then) even if it is large today. In the last decade interest in such theories has been revitalized as a result of early Universe studies. Extended inflation explicitly requires a model in which G varies [78] and, in higher dimensional Kaluza-Klein-type cosmologies, the variation in the sizes of the extra dimensions also naturally leads to this [39,74,88].

The behaviour of homogeneous cosmological models in BD theory is well understood [6]. They are vacuum-dominated at early times but always tend towards the general relativistic solution during the radiation-dominated era. This means that the full radiation solution can be approximated by joining a BD vacuum solution to a general relativistic radiation solution at some time which may be regarded as a free parameter of the theory. However, when the matter density becomes greater than the radiation density at around 10^5y, the equation of state becomes dustlike ($p = 0$) and G begins to vary again.

The consequences of the cosmological variation of G for PBH evaporation depend upon how the value of G near the black hole evolves. Barrow [4] introduces two possibilities: in scenario A, G everywhere maintains the background cosmological value (so ϕ is homogeneous); in scenario B, it preserves the value it had at the formation epoch near the black hole even though it evolves at large distances (so ϕ becomes inhomogeneous). On the assumption that a PBH of mass M has a temperature and mass-loss rate

$$T \propto (GM)^{-1}, \quad \dot{M} \propto (GM)^{-2}, \tag{15}$$

with $G = G(t)$ in scenario A and $G = G(M)$ in scenario B, Barrow & Carr [5] calculate how the evaporation constraints summarized in Fig. 1 are modified for a wide range of varying-G models. The question of whether scenario A or scenario B is more plausible has been studied in several papers [21,43,52,65] but is still unresolved.

6 PBHs as a Probe of Gravitational Collapse

The criterion for PBH formation given in Sect. 3 is rather simplistic and not based on a detailed calculation. The first numerical studies of PBH formation were carried out by Nadezhin et al. [92]. These roughly confirmed the criterion $\delta > \gamma$ for PBH formation, although the PBHs could be somewhat smaller than the horizon. In recent years several groups have carried out more detailed hydrodynamical calculations and these have refined the $\delta > \gamma$ criterion and hence the estimate for $\beta(M)$ given by (4). Niemeyer & Jedamzik [96] find that one needs $\delta > 0.8$ rather than $\delta > 0.3$ to ensure PBH formation and they also find that there is little accretion after PBH formation, as expected theoretically [22]. Shibata & Sasaki [108] reach similar conclusions.

A particularly interesting development has been the application of "critical phenomena" to PBH formation. Studies of the collapse of various types of spherically symmetric matter fields have shown that there is always a critical solution which separates those configurations which form a black hole from those which disperse to an asymptotically flat state. The configurations are described by some index p and, as the critical index p_c is approached, the black hole mass is found to scale as $(p - p_c)^\eta$ for some exponent η. This effect was first discovered for scalar fields [26] but subsequently demonstrated for radiation [35] and then more general fluids with equation of state $p = \gamma\rho$ [75,90].

In all these studies the spacetime was assumed to be asymptotically flat. However, Niemeyer & Jedamzik [95] have recently applied the same idea to study black hole formation in asymptotically Friedmann models and have found similar results. For a variety of initial density perturbation profiles, they find that the relationship between the PBH mass and the horizon-scale density perturbation has the form

$$M = K M_{\mathrm{H}} (\delta - \delta_c)^\gamma \,, \tag{16}$$

where M_{H} is the horizon mass and the constants are in the range $0.34 < \gamma < 0.37$, $2.4 < K < 11.9$ and $0.67 < \delta_c < 0.71$ for the various configurations. Since $M \to 0$ as $\delta \to \delta_c$, this suggests that PBHs may be much smaller than the particle horizon at formation and it also modifies the mass spectrum [45,48,76,117]. However, it is clear that a fluid description must break down if they are too small and recent calculations by Hawke & Stewart [53] show that black holes can only form on scales down to 10^{-4} of the horizon mass.

There has also been interest recently in whether PBHs could have formed at the quark-hadron phase transition at 10^{-5}s because of a temporary softening of the equation of state then. Such PBHs would naturally have the sort of mass required to explain the MACHO microlensing results [66]. If the QCD phase transition is assumed to be of 1st order, then hydrodynamical calculations show that the value of δ required for PBH formation is indeed reduced below the value which pertains in the radiation case [67]. This means that PBH formation will be strongly enhanced at the QCD epoch, with the mass distribution being peaked around the horizon mass. One of the interesting implications of this scenario is the possible existence of a halo population of *binary* black holes [93]. With a full halo of such objects, there could then be 10^8 binaries inside 50 kpc and some of these could be coalescing due to gravitational radiation losses at the present epoch. If the associated gravitational waves were detected, it would provide a unique probe of the halo distribution [62].

7 PBHs as a Probe of High Energy Physics

We have seen that a black hole of mass M will emit particles like a black-body of temperature [56]

$$T \approx 10^{26} \left(\frac{M}{\mathrm{g}}\right)^{-1} \mathrm{K} \approx \left(\frac{M}{10^{13}\ \mathrm{g}}\right)^{-1} \mathrm{GeV}. \tag{17}$$

This assumes that the hole has no charge or angular momentum. This is a reasonable assumption since charge and angular momentum will also be lost through quantum emission but on a shorter timescale than the mass [100]. This means that it loses mass at a rate

$$\dot{M} = -5 \times 10^{25} \left(\frac{M}{\mathrm{g}}\right)^{-2} f(M)\ \mathrm{g\ s}^{-1} \tag{18}$$

where the factor $f(M)$ depends on the number of particle species which are light enough to be emitted by a hole of mass M, so the lifetime is

$$\tau(M) = 6 \times 10^{-27} f(M)^{-1} \left(\frac{M}{\mathrm{g}}\right)^{3}\ \mathrm{s}. \tag{19}$$

The factor f is normalized to be 1 for holes larger than 10^{17} g and such holes are only able to emit "massless" particles like photons, neutrinos and gravitons. Holes in the mass range $10^{15}\ \mathrm{g} < M < 10^{17}$ g are also able to emit electrons, while those in the range $10^{14}\ \mathrm{g} < M < 10^{15}$ g emit muons which subsequently decay into electrons and neutrinos. The latter range includes, in particular, the critical mass for which τ equals the age of the Universe. If the total density parameter is 1, this can be shown to be $M_* = 4.4 \times 10^{14} h^{-0.3}$g where h is the Hubble parameter in units of 100 [84].

Once M falls below 10^{14}g, a black hole can also begin to emit hadrons. However, hadrons are composite particles made up of quarks held together by gluons. For temperatures exceeding the QCD confinement scale of $\Lambda_{\mathrm{QCD}} = 250$–300 GeV, one would therefore expect these fundamental particles to be emitted rather than composite particles. Only pions would be light enough to be emitted below Λ_{QCD}. Since there are 12 quark degrees of freedom per flavour and 16 gluon degrees of freedom, one would also expect the emission rate (i.e. the value of f) to increase dramatically once the QCD temperature is reached.

The physics of quark and gluon emission from black holes is simplified by a number of factors. Firstly, one can show that the separation between successively

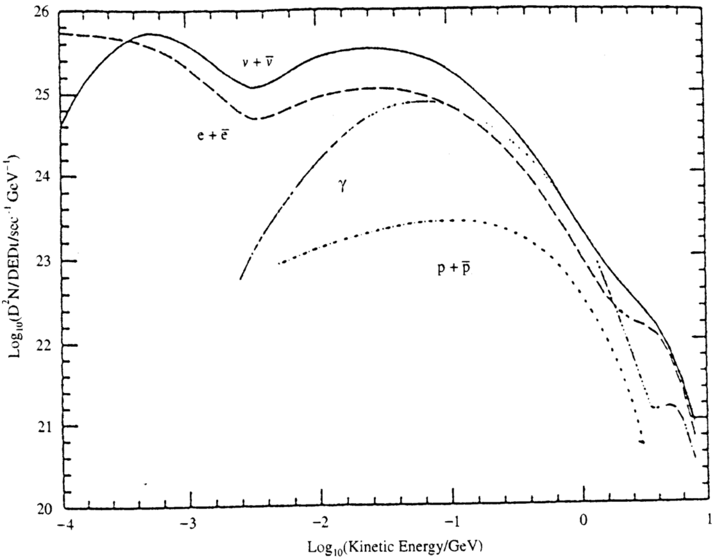


Fig. 4. Instantaneous emission from a 1 GeV black hole. Plotted is the number of particles emitted per time for energy interval in appropriate units.

emitted particles is about 20 times their wavelength, which means that short range interactions between them can be neglected. Secondly, the condition $T > \Lambda_{\mathrm{QCD}}$ implies that their separation is much less than $\Lambda_{\mathrm{QCD}}^{-1} \approx 10^{-13}$cm (the characteristic strong interaction range) and this means that the particles are also unaffected by strong interactions. The implication of these three conditions is that one can regard the black hole as emitting quark and gluon jets of the kind produced in collider events. The jets will decay into hadrons over a distance which is always much larger than GM, so gravitational effects can be neglected. The hadrons may then decay into astrophysically stable particles through weak and electomagnetic decays.

To find the final spectra of stable particles emitted from a black hole, one must convolve the Hawking emission spectrum with the jet fragmentation function. This gives the instantaneous emission spectrum shown in Fig. 4 for a $T = 1$ GeV black hole [85]. The direct emission just corresponds to the small bumps on the right. All the particle spectra show a peak at 100 MeV due to pion decays; the electrons and neutrinos also have peaks at 1 MeV due to neutron decays. In order to determine the present day background spectrum of particles generated by PBH evaporations, one must first integrate over the lifetime of each hole of mass M and then over the PBH mass spectrum [85]. In doing this, one must allow for the fact that smaller holes will evaporate at an earlier cosmological epoch, so the particles they generate will be redshifted in energy by the present epoch.

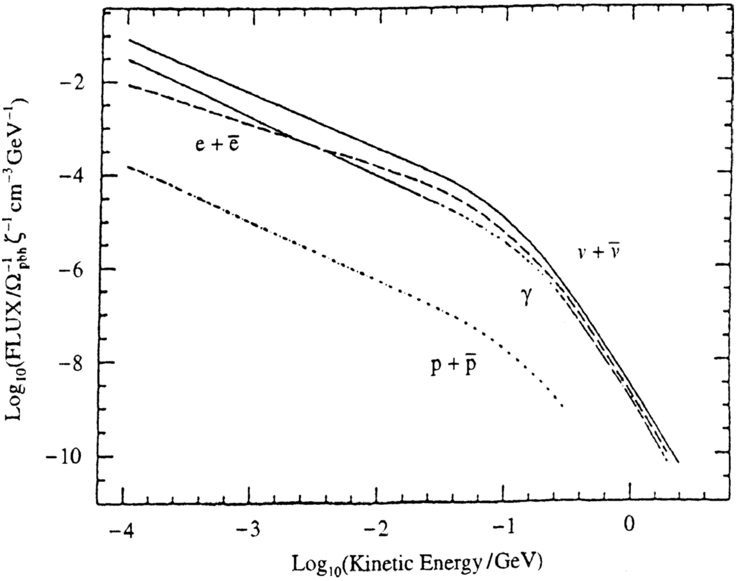


Fig. 5. Spectrum of particles from uniformly distributed PBHs

If the holes are uniformly distributed throughout the Universe, the background spectra should have the form indicated in Fig. 5. All the spectra have rather similar shapes: an E^{-3} fall-off for $E > 100$ MeV due to the final phases of evaporation at the present epoch and an E^{-1} tail for $E < 100$ MeV due to the fragmentation of jets produced at the present and earlier epochs. Note that the E^{-1} tail generally masks any effect associated with the mass spectrum of smaller PBHs which evaporated at earlier epochs [19].

The situation is more complicated if the PBHs evaporating at the present epoch are clustered inside our own Galactic halo (as is most likely). In this case, any charged particles emitted after the epoch of galaxy formation (i.e. from PBHs only somewhat smaller than M_*) will have their flux enhanced relative to the photon spectra by a factor ξ which depends upon the halo concentration factor and the time for which particles are trapped inside the halo by the Galactic magnetic field. This time is rather uncertain and also energy-dependent. At 100 MeV one has $\xi \sim 10^3$ for electrons or positrons and $\xi \sim 10^4$ for protons and antiprotons. MacGibbon & Carr [84] first used the observed cosmic ray spectra to constrain $\Omega_{\rm PBH}$ but their estimates have recently been updated.

7.1 Gamma-Rays

Recent EGRET observations [109] give a γ-ray background of

$$\frac{dF_\gamma}{dE} = 7.3(\pm 0.7) \times 10^{-14} \left(\frac{E}{100\ \mathrm{MeV}}\right)^{-2.10\pm 0.03} \mathrm{cm}^{-3}\mathrm{GeV}^{-1} \qquad (20)$$

between 30 MeV and 120 GeV. Carr & MacGibbon [24] showed that this leads to an upper limit

$$\Omega_{\rm PBH} \leq (5.1 \pm 1.3) \times 10^{-9} h^{-2}, \tag{21}$$

which is a refinement of the original Page-Hawking limit, but the form of the spectrum suggests that PBHs do not provide the dominant contribution. If PBHs are clustered inside our own Galactic halo, then there should also be a Galactic γ-ray background and, since this would be anisotropic, it should be separable from the extragalactic background. The ratio of the anisotropic to isotropic intensity depends on the Galactic longtitude and latitude, the ratio of the core radius to our Galactocentric radius, and the halo flattening. Wright claims that such a halo background has been detected [114]. His detailed fit to the EGRET data, subtracting various other known components, requires the PBH clustering factor to be $(2-12) \times 10^5 h^{-1}$, comparable to that expected.

7.2 Antiprotons

Since the ratio of antiprotons to protons in cosmic rays is less than 10^{-4} over the energy range 100 MeV $-$ 10 GeV, whereas PBHs should produce them in equal numbers, PBHs could only contribute appreciably to the antiprotons [111]. It is usually assumed that the observed antiproton cosmic rays are secondary particles, produced by spallation of the interstellar medium by primary cosmic rays. However, the spectrum of secondary antiprotons should show a steep cut-off at kinetic energies below 2 GeV, whereas the spectrum of PBH antiprotons should increase with decreasing energy down to 0.2 GeV, so this provides a distinct signature [73].

MacGibbon & Carr originally calculated the PBH density required to explain the interstellar antiproton flux at 1 GeV and found a value somewhat larger than the γ-ray limit [84]. More recent data on the antiproton flux below 0.5 GeV comes from the BESS balloon experiment [118] and Maki et al. [89] have tried to fit this data in the PBH scenario. They model the Galaxy as a cylindrical diffusing halo of diameter 40 kpc and thickness 4-8 kpc and then using Monte Carlo simulations of cosmic ray propagation. A comparison with the data shows no positive evidence for PBHs (i.e. there is no tendency for the antiproton fraction to tend to 0.5 at low energies) but they require the fraction of the local halo density in PBHs to be less than 3×10^{-8} and this is stronger than the γ-ray background limit. A more recent attempt to fit the observed antiproton spectrum with PBH emission comes from Barrau et al. [8] and is shown in Fig. 6. A key test of the PBH hypothesis will arise during the solar minimum period because the flux of primary antiprotons should be enhanced then, while that of the secondary antiprotons should be little affected [91].

7.3 PBH Explosions

One of the most striking observational consequences of PBH evaporations would be their final explosive phase. However, in the standard particle physics picture,

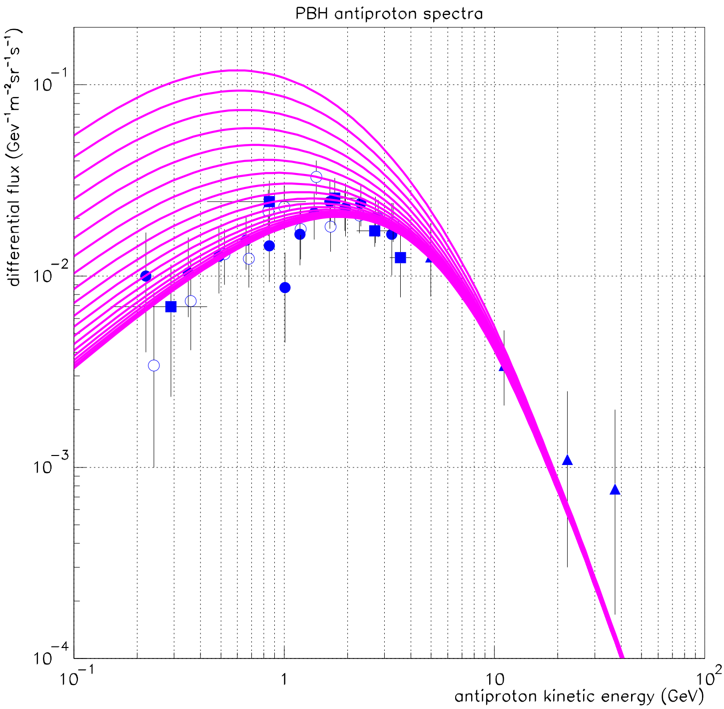


Fig. 6. Comparison of PBH emission and antiproton data from Barrau et al.

where the number of elementary particle species never exceeds around 100, the likelihood of detecting such explosions is very low. Indeed, in this case, observations only place an upper limit on the explosion rate of $5 \times 10^8 \mathrm{pc}^{-3}\mathrm{y}^{-1}$ [1,107]. This compares to Wright's γ-ray halo limit of 0.3 $\mathrm{pc}^{-3}\mathrm{y}^{-1}$ and the Maki et al. antiproton limit of 0.02 $\mathrm{pc}^{-3}\mathrm{y}^{-1}$.

However, the physics at the QCD phase transition is still uncertain and the prospects of detecting explosions would be improved in less conventional particle physics models. For example, in a Hagedorn-type picture, where the number of particle species exponentiates at the quark-hadron temperature, the upper limit is reduced to 0.05 $\mathrm{pc}^{-3}\mathrm{y}^{-1}$ [37]. Cline and colleagues have argued that one might expect the formation of a QCD fireball at this temperature [27] and this might even explain some of the short period γ-ray bursts observed by BATSE [28]. They claim to have found 42 candidates of this kind and the fact that their distribution matches the spiral arms suggests that they are Galactic. Although this proposal is speculative and has been disputed [46], it has the attraction of making testable predictions (eg. the hardness ratio should increase as the duration of the burst decreases). A rather different way of producing a γ-ray burst is to assume that the outgoing charged particles form a plasma due to turbulent magnetic field effects at sufficiently high temperatures [10].

Some people have emphasized the possibility of detecting very high energy cosmic rays from PBHs using air shower techniques [31,51,77]. However, recently these efforts have been set back by the claim of Heckler [59] that QED interactions could produce an optically thick photosphere once the black hole temperature exceeds $T_{\rm crit} = 45$ GeV. In this case, the mean photon energy is reduced to $m_e(T_{\rm BH}/T_{\rm crit})^{1/2}$, which is well below $T_{\rm BH}$, so the number of high energy photons is much reduced. He has proposed that a similar effect may operate at even lower temperatures due to QCD effects [60]. Several groups have examined the implications of this proposal for PBH emission [29,69]. However, these arguments should not be regarded as definitive since MacGibbon et al. claim that QED and QCD interactions are never important [87].

8 PBHs as a Probe of Quantum Gravity

In the standard Kaluza-Klein picture, the extra dimensions are assumed to be compactified on the scale of the Planck length. This means that the influence of these extra dimensions only becomes important at an energy scale of 10^{19}GeV and this is also presumably the scale on which quantum gravity effects become significant. In particular, such effects are only important for black hole evaporations once the black hole mass gets down to the Planck mass of 10^{-5}g. Conceivably, this could result in black hole evaporation ceasing, so that one ends up with stable Planck-mass relics, and this leads to the sort of "relics" constraints indicated in Figs. 1, 2, and 3. Various non-quantum-gravitational effects (such as higher order corrections to the gravitational Lagrangian or string effects) could also lead to stable relics [25] but the relic mass is always close to the Planck mass.

In "brane" versions of Kaluza-Klein theory, some of the extra dimensions can be much larger than the Planck length and this means that quantum gravity effects may become important at a much smaller energy scale. If the internal space has n dimensions and a compact volume V_n, then Newton's constant G_N is related to the higher dimensional gravitational constant G_D and the value of the modified Planck mass $M_{\rm Pl}$ is related to the usual 4-dimensional Planck mass M_4 by the order-of-magnitude equations:

$$G_N \sim \frac{G_D}{V_n}, \qquad M_{\rm Pl}^{n+2} \sim \frac{M_4^2}{V_n}. \tag{22}$$

The same relationship applies if one has an infinite extra dimension but with a "warped" geometry, provided one interprets V_n as the "warped volume". In the standard model, $V_n \sim 1/M_4^n$ and so $M_{\rm Pl} \sim M_4$. However, with large extra dimensions, one has $V_n >> 1/M_4^n$ and so $M_{\rm Pl} << M_4$. In particular, this might permit quantum gravitational effects to arise at the experimentally observable TeV scale.

If this were true, it would have profound implications for black hole formation and evaporation since black holes could be generated in accelerator experiments, such as the Large Hadron Collider (LHC). Two partons with centre-of-mass energy $\sqrt{s}$ will form a black hole if they come within a distance corresponding to the Schwarzschild radius r_S for a black hole whose mass $M_{\rm BH}$ is equivalent to that energy [33,42,106]. Thus the cross-section for black hole production is

$$\sigma_{BH} \approx \pi r_S^2 \Theta(\sqrt{s} - M_{\rm BH}^{\rm min}) \,, \tag{23}$$

where $M_{\rm BH}^{\rm min}$ is the mass below which the semi-classical approximation fails. Here the Schwarzschild radius itself depends upon the number of internal dimensions:

$$r_S \approx \frac{1}{M_{\rm Pl}} \left(\frac{M_{\rm BH}}{M_{\rm Pl}} \right)^{1/(1+n)} \,, \tag{24}$$

so that $\sigma_{BH} \propto s^{1/(n+1)}$. This means that the cross-section for black hole production in scattering experiments goes well above the cross-section for the standard model above a certain energy scale and in a way which depends on the number of extra dimensions.

The evaporation of the black holes produced in this way will produce a characteristic signature [33,42,106] because the temperature and lifetime of the black holes depend on the number of internal dimensions:

$$T_{BH} \approx \frac{n+1}{r_{\rm S}} \,, \qquad \tau_{\rm BH} \approx \frac{1}{M_{\rm Pl}} \left(\frac{M_{\rm BH}}{M_{\rm Pl}} \right)^{(n+3)/(n+1)} \,. \tag{25}$$

Thus the temperature is decreased relative to the standard 4-dimensional case and the lifetime is increased. The important qualitative effect is that a large fraction of the beam energy is converted into transverse energy, leading to large-multiplicity events with many more hard jets and leptons than would otherwise be expected. In principle, the formation and evaporation of black holes might be observed by LHC by the end of the decade and this might also allow one to experimentally probe the number of extra dimensions. On the other hand, this would also mean that scattering processes above the Planck scale could not be probed directly because they would be hidden behind a black hole event horizon.

Similar effects could be evident in the interaction between high energy cosmic rays and atmospheric nucleons. Nearly horizontal cosmic ray neutrinos would lead to the production of black holes, whose decays could generate deeply penetrating showers with an electromagnetic component substantially larger than that expected with conventional neutrino interactions. Several authors have studied this in the context of the Pierre Auger experiment, with event rates in excess of one per year being predicted [2,36,106]. Indeed there is a small window of opportunity in which Auger might detect such events before LMC.

It should be stressed that the black holes produced in these processes should not themselves be described as "primordial" since they do not form in the early Universe. On the other hand, it is clear that the theories which predict such processes will also have profound implications for the formation and evaporation of those black holes which *do* form then. This is because, at sufficiently early times, the effects of the extra dimensions must be cosmologically important. However, these effects are not yet fully understood.

9 Conclusions

We have seen that PBHs could provide a unique probe of the early Universe, gravitational collapse, high energy physics and quantum gravity. In the "early Universe" context, particularly useful constraints can be placed on inflationary scenarios and on models in which the value of the gravitational "constant" G varies with cosmological epoch. In the "gravitational collapse" context, the existence of PBHs could provide a unique test of the sort of critical phenomena discovered in recent numerical calculations. In the "high energy physics" context, information may come from observing cosmic rays from evaporating PBHs since the constraints on the number of evaporating PBHs imposed by gamma-ray background observations do not exclude their making a significant contribution to the Galactic flux of electrons, positrons and antiprotons. Evaporating PBHs may also be detectable in their final explosive phase as gamma-ray bursts if suitable physics is invoked at the QCD phase transition. In the "quantum gravity" context, the formation and evaporation of small black holes could lead to observable signatures in cosmic ray events and accelerator experiments, provided there are extra dimensions and provided the quantum gravity scale is around a TeV.

References

1. D.E. Alexandreas et al.: *Phys. Rev. Lett.* **71**, 2524 (1993)
2. L. Anchordogui and H. Goldberg: *Phys. Rev. D* **65**, 047502 (2002)
3. J.D. Barrow: *MNRAS* **192**, 427 (1980)
4. J.D. Barrow: *Phys. Rev. D* **46**, R3227 (1992)
5. J.D. Barrow and B.J. Carr: *Phys. Rev. D* **54** (1996) 3920
6. J.D. Barrow and P. Parsons: *Phys. Rev. D* **55**, 1906 (1997)
7. J.D. Barrow, E.J. Copeland and A.R. Liddle: *Phys. Rev. D* **46**, 645 (1992)
8. A. Barrau et al.: Astron.Astrophys., in press (2002); astro-ph/0112486
9. B.A. Bassett and S. Tsujikawa: *Phys. Rev. D* **63**, 123503 (2001)
10. A.A. Belyanin et al.: (1997). Preprint (unpublished)
11. P.G. Bergmann: *Int. J. Theor. Phys.* **1**, 25 (1968)
12. G.V. Bicknell and R.N. Henriksen: *Ap. J.* **219**, 1043 (1978)
13. M.J. Bowick et al.: *Phys. Rev. Lett.* **61**, 2823 (1988)
14. C. Brans and R.H. Dicke: *Phys. Rev.* **124**, 925 (1961)
15. T. Bringmann, C. Kiefer and D. Polarsk:, *Phys. Rev. D* **65**, 024008 (2002)
16. J.S. Bullock and J.R. Primack: *Phys. Rev. D* **55**, 7423 (1997)
17. R. Caldwell and P. Casper: *Phys. Rev. D* **53**, 3002 (1996)
18. B.J. Carr: *Ap. J.* **201**, 1 (1975)
19. B.J. Carr: *Ap. J.* **206**, 8 (1976)
20. B.J. Carr: in *Observational and Theoretical Aspects of Relativistic Astrophysics and Cosmology*, ed. J.L. Sanz and L.J. Goicoechea (World Scientific, Singapore, 1985), p. 1
21. B.J. Carr and C.A. Goymer: *Prog. Theor. Phys.* **136**, 321 (1999)
22. B.J. Carr and S.W. Hawking: *MNRAS* **168**, 399 (1974)
23. B.J. Carr and J.E. Lidsey: *Phys. Rev. D* **48**, 543 (1993)

24. B.J. Carr and J.H. MacGibbon: *Phys. Rep.* **307**, 141 (1998)
25. B.J. Carr, J.H. Gilbert and J.E. Lidsey: *Phys. Rev. D* **50**, 4853 (1994)
26. M.W. Choptuik: *Phys. Rev. Lett.* **70**, 9 1993)
27. D.B. Cline and W. Hong: *Ap. J. Lett.* **401**, L57 (1992)
28. D.B.Cline, D.A. Sanders and W. Hong: *Ap. J.* **486**, 169 (1997)
29. J. Cline, M. Mostoslavsky and G. Servant: *Phys. Rev. D* **59**, 063009 (1999)
30. S. Coleman, J. Preskill and F. Wilczek: *Mod. Phys. Lett. A* **6**, 1631 (1991)
31. D.G. Coyne, C. Sinnis and R. Somerville: in *Proceedings of the Houston Advanced Research Center Conference on Black Holes* (1992)
32. M. Crawford and D.N. Schramm: *Nature* **298**, 538 (1982)
33. S. Dimopoulos and G. Landsberg: *Phys. Rev. Lett.* **87**, 161602 (2001)
34. R. Easther and M. Parry: *Phys. Rev. D* **62**, 103503 (2000)
35. C.R. Evans and J.S. Coleman: *Phys. Rev. Lett.* **72**, 1782 (1994)
36. J.L. Feng and A.D. Shapere: *Phys. Rev. Lett.* **88**, 021303 (2002)
37. C.E. Fichtel et al.: *Ap. J.* **1434**, 557 (1994)
38. F. Finelli and S. Khlebnikov: *Phys. Lett. B* **504**, 309 (2001)
39. P.G.O. Freund: *Nuc. Phys. B* **209**, 146 (1982)
40. J. Garcia-Bellido, A. Linde and D. Wands: *Phys. Rev. D* **54**, 6040 (1997)
41. J. Garriga and M. Sakellariadou: *Phys. Rev. D* **48**, 2502 (1993)
42. S.B. Giddings and S. Thomas: *Phys. Rev. D* **65**, 056010 (2002)
43. C. Goymer and B.J. Carr, unpublished, preprint (1999)
44. A.M. Green: *Phys. Rev. D* **60**, 063516 (1999)
45. A.M. Greenv *Ap. J.* **537**, 708 (2000)
46. A.M. Green: *Phys. Rev. D* **65**, 027301 (2002)
47. A.M. Green and A.R. Liddle: *Phys. Rev. D* **56**, 6166 (1997)
48. A.M. Green and A.R. Liddle: *Phys. Rev. D* **60**, 063509 (1999)
49. A.M. Green, A.R. Liddle and A. Riotto: *Phys. Rev. D* **56**, 7559 (1997)
50. A.M. Green and K.A. Malik: *Phys. Rev. D* **64**, 021301 (2001)
51. F. Halzen, E. Zas, J. MacGibbon and T.C. Weekes: *Nature* **298**, 538 (1991)
52. T. Harada, B.J. Carr and C.A. Goymer: *Phys. Rev. D* **66**, 104023 (2002)
53. I. Hawke and J.M. Stewart: *Class. Quantum Grav.* **19**, 3687 (2002)
54. S.W. Hawking: *MNRAS* **152**, 75 (1971)
55. S.W. Hawking: *Nature* **248**, 30 (1974)
56. S.W. Hawking: *Comm. Math. Phys.* **43**, 199 (1975)
57. S.W. Hawking: *Phys. Lett. B* **231**, 237 (1989)
58. S.W. Hawking: I. Moss and J. Stewart, *Phys. Rev. D* **26**, 2681 (1982)
59. A. Heckler: *Phys. Rev. D* **55**, 840 (1997)
60. A. Heckler: *Phys. Lett. B* **231**, 3430 (1997)
61. H.M. Hodges and G.R. Blumenthal: *Phys. Rev. D* **42**, 3329 (1990)
62. K. Ioka, T. Tanaka and T. Nakamura et al.: *Phys. Rev. D* **60**, 083512 (1999)
63. P. Ivanov: *Phys. Rev. D* **57**, 7145 (1998)
64. P. Ivanov, P. Naselsky and I. Novikov: *Phys. Rev. D* **50**, 7173 (1994)
65. T. Jacobsen: *Phys. Rev. Lett.* **83**, 2699 (1999)
66. K. Jedamzik: *Phys. Rev. D* **55**, R5871 (1997); *Phys. Rep.* **307**, 155 (1998)
67. K. Jedamzik and J. Niemeyer: *Phys. Rev. D* **59**, 124014 (1999)
68. T. Kanazawa, M. Kawasaki and T. Yanagida: *Phys. Lett. B* **482**, 174 (2000)
69. J. Kapusta: in *Phase Transitions in the Early Universe: Theory and Observations*, ed. H.J. de Vega et al. (Kluwer 2001), p. 471
70. M.Yu. Khlopov and A.G. Polnarev: *Phys. Lett. B* **97**, 383 (1980)
71. H. Kim: *Phys. Rev. D* **62**, 063504 (2000)

72. H. Kim: C.H. Lee and J.H. MacGibbon, *Phys. Rev. D* **59**, 063004 (1999)
73. P. Kiraly et al.: *Nature* **293**, 120 (1981)
74. E.W. Kolb, M.J. Perry and T.P. Walker: *Phys. Rev. D* **33**, 869 (1986)
75. T. Koike, T. Hara and S. Adachi: *Phys. Rev. D* **59**, 104008 (1999)
76. G.D. Kribs, A.K. Leibovich and I.Z. Rothstein: *Phys. Rev. D* **60**, 103510 (1999)
77. F. Krennrich, S. Le Bohec and T.C. Weekes: *Ap. J.* **529**, 506 (2000)
78. D. La and P.J.Steinhardt: *Phys. Lett. B* **220**, 375 (1989)
79. S.M. Leach, I.J. Grivell and A.R. Liddle: *Phys. Rev. D* **62**, 043516 (2000)
80. D.N.C. Lin, B.J. Carr and S.M. Fall: *MNRAS* **177**, 51 (1976)
81. J.E. Lidsey, T. Matos and L.A. Urena-Lopez: *Phys. Rev. D* **66**, 023514 (2002)
82. J.H. MacGibbon: *Nature* **329**, 308 (1987)
83. J.H. MacGibbon: *Phys. Rev. D* **44**, 376 (1991)
84. J.H. MacGibbon and B.J. Carr: *Ap. J.* **371**, 447 (1991)
85. J.H. MacGibbon and B.R. Webber: *Phys. Rev. D* **41**, 3052 (1990)
86. J.H. MacGibbon, R.H. Brandenberger and U.F. Wichoski: *Phys. Rev. D* **57**, 2158 (1998)
87. J.H. MacGibbon, B.J. Carr and D.N. Page: unpublished, preprint (2002)
88. K. Maeda: *Class. Quant. Grav.* **3**, 233 (1986)
89. K. Maki, T. Mitsui and S. Orito: *Phys. Rev. Lett.* **76**, 3474 (1996)
90. D. Maison: *Phys. Lett. B* **366**, 82 (1996)
91. I.V Mosalenko et al.: in *The Outer Heliosphere: The Next Frontier*, eds. H.J. Fahr et al. (2001)
92. D.K. Nadezhin, I.D. Novikov and A.G. Polnarev: *Sov. Astron.* **22**, 129 (1978)
93. T. Nakamura, M. Sasaki, T. Tanaka and K. Thorne: *Ap. J.* **487**, L139 (1997)
94. P.D. Naselsky & A.G. Polnarev: *Sov. Astron.* **29**, 487 (1985)
95. J. Niemeyer and K. Jedamzik: *Phys. Rev. Lett.* **80**, 5481 (1998)
96. J. Niemeyer and K. Jedamzik: *Phys. Rev. D* **59**, 124013 (1999)
97. K. Nordtvedt: *Ap. J.* **161**, 1059 (1970)
98. I.D. Novikov, A.G. Polnarev, A.A. Starobinsky and Ya.B. Zeldovich: *Astron. Astrophys.* **80**, 104 (1979)
99. P.N. Okeke and M.J. Rees: *Astron. Astrophys.* **81**, 263 (1980)
100. D.N. Page: *Phys. Rev. D* **16**, 2402 (1977)
101. D.N. Page and S.W. Hawking: *Ap. J.* **206**, 1 (1976)
102. A.G. Polnarev and R. Zemboricz: *Phys. Rev. D* **43**, 1106 (1988)
103. N.A. Porter and T.C. Weekes: *Nature* **277**, 199 (1979)
104. L. Randall, M. Soljacic and A.H. Guth: *Nucl. Phys. B* **472**, 377 (1996)
105. M.J. Rees: *Nature* **266**, 333 (1977)
106. A. Ringwald and H. Tu: *Phys. Lett. B* **525**, 135 (2002)
107. D.V. Semikoz: *Ap. J.* **436**, 254 (1994)
108. M. Shibata and M. Sasaki: *Phys. Rev. D* **60**, 084002 (1999)
109. P. Sreekumar et al.: *Ap. J.* **494**, 523 (1998)
110. A. Taruya: *Phys. Rev. D* **59**, 103505 (1999)
111. M.S. Turner: *Nature* **297**, 379 (1982)
112. R.V. Wagoner: *Phys. Rev. D* **1**, 3209 (1970)
113. C.M. Will: *Theory and Experiment in Gravitational Physics* (Cambridge University Press, Cambridge, 1993)
114. E.L. Wright: *Ap. J.* **459**, 487 (1996)
115. J. Yokoyama: *Astron. Astrophys.* **318**, 673 (1997)
116. J. Yokoyama: *Phys. Rev. D* **58**, 083510 (1998)
117. J. Yokoyama: *Phys. Rev. D* **58**, 107502 (1998)
118. K. Yoshimura et al.: *Phys. Rev. Lett.* **75**, 3792 (1995)
119. Ya.B. Zeldovich and I.D. Novikov: *Sov. Astron. A. J.* **10**, 602 (1967)
120. Ya.B. Zeldovich and A.A. Starobinsky: *JETP Lett.* **24**, 571 (1976)

On the Assignment of Entropy to Black Holes

Daniel Sudarsky

Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México,
A. Postal 70-543, México D.F. 04510, México

Abstract. The range of validity of the usual identification of the black hole entropy with the area of the horizon is considered from a general point of view. The situation is then revised in the light of an example in which the actual presence of the event horizon on a given hypersurface depends on a quantum event which occurs in the future of the given hypersurface. This situation indicates that there is something fundamental that is missing in our current ideas about the nature of a theory of quantum gravity, or, alternatively, that there is something fundamental that we do not understand about entropy in general, or at least in its association with black holes.

1 Introduction

It is well known by now that the laws of black hole mechanics seem to indicate a deep connection between gravitation, thermodynamics and quantum mechanics [1]. In this regard the black hole entropy is thought to play a central role regarding possible clues about the nature of the quantum theory of gravitation. The laws of Black Hole Mechanics together with Hawking's discovery [2] of the thermal radiation, of quantum mechanical origin, by black holes has lead to the identification of the black hole area and its entropy.

We must keep in mind the difference in the two levels at which entropy can be analyzed, the thermodynamical and the statistical mechanical. At the thermodynamical level, one does not connect the microscopic description of the system with its macroscopic one, while in the statistical mechanical approach one can in principle evaluate directly the quantities like the entropy of the system starting from the theory of the relevant microscopic degrees of freedom. For instance, starting from the Hamiltonian of a set of non-interacting elementary "atoms" one finds, trough a systematic calculation, all thermodynamic functions for an ideal gas. Thus, the evaluation of the entropy of a black hole by statistical mechanical methods and a comparison with the " known" value, has for a long time been considered as a fundamental test of proposed theories of quantum gravity, and as a key method to distinguishing between them.

It is then to a certain degree a disappointment (as far as the usefulness of the test to distinguish between theories is concerned) that currently the two most popular approaches towards a theory of quantum gravity, the String Theory approach [3] and the Loop Quantum Gravity approach [4], have, each within a somehow restricted domain, achieved success in recovering the $S = A/4$ result (in this work we will use units in which $G = c = \hbar = k_B = 1$).

D. Sudarsky, On the Assignment of Entropy to Black Holes, Lect. Notes Phys. **631**, 323–334 (2003)
http://www.springerlink.com/

It is thus natural to consider more complex situations and see how the different approaches manage the challenge. This is certainly a path that many researchers in the area are eagerly exploring. Here we will consider a situation which is in some sense an extreme manifestation of this approach, and which seems to indicate that there is either something fundamentally wrong with a large class of approaches towards a theory of quantum gravity, including both the String Theory approach and the Loop Quantum Gravity approach, or there is something fundamental about entropy, and in particular, the assignment of entropy to black holes that is beyond our current understanding. The discussion here is mostly based on previous work carried out in collaboration with R. Sorkin [5] and A. Corichi [6].

2 The Assignment of Entropy

Let us start by considering the issue in a general setting: Under what conditions is the assignment of entropy appropriate, i.e. *when?, to what?, and based on what data?* do we expect to assign an entropy? (See [7] for a general analysis of the conceptual issues). Let us concentrate, unless otherwise specified, on the statistical mechanical entropy $S^{\rm stat}$, as opposed to the thermodynamical entropy $S^{\rm thermo}$, as the first is obviously more fundamental than the second. Next one can address the issue of whether it is the Gibbs or the Boltzman entropy what one should be considering. Recall that Gibbs entropy is associated with an ensemble of systems, while the Boltzman Entropy is associated with a single system whose state is specified only up to a certain limited precision that defines the "mesostate". In practice, both rely in the introduction of a coarse graining defining a certain volume in phase space, or the distribution of the ensemble over such volume, of which entropy is a measure. Thus we will not further specify the particular notion of $S^{\rm stat}$ we are talking about.

The main points we want to make here are the following:

(i) *Entropy should be assigned in every situation*, not only to equilibrium, (or quasi-equilibrium) situations. Otherwise the second law would lose much of its predictive power: One would not be able to argue, for instance that an isolated system tends towards equilibrium as a result of the tendency of entropy to increase, if entropy is not defined in the intermediate, non-equilibrium situations. Furthermore, one would not be able to rule out perpetual motion machines, simply on the grounds of this law, since one could certainly propose a machine that avoids passing trough an equilibrium situation, and thus never having an entropy associated to it. Finally, one expects the laws of physics to have a Markovian nature, in the sense that their predicting power should not increase with the knowledge of the past together with the present as opposed to the knowledge of the present alone. If it were the case that entropy was defined only in certain situations, and at the present time a certain (isolated) system did not have an entropy defined, we would not be able to rule out its possible evolution towards a situation where its entropy would be S_0, but if we knew that in the past its entropy was $S_1 > S_0$ we would be able to rule out such evolution. It is very unlikely that other physical laws would restore the "Markovianness" of physics because, as far as we know, the second law is the only law of nature containing an explicit arrow of time (with the exception of some details of the weak interaction that are not expected to play a role in the issue at hand).

The arguments above also indicate that:

(ii) *the entropy should be associated with an instantaneous situation (or, more generally, with a situation localized in time)*, rather than, say, with the full world path (or history) of a system, for otherwise the predictive power of the second law will disappear. In the general relativistic context this means that entropy should be associated with a space-like hypersurface.

Moreover, as already mentioned, entropy should be assigned to a state of the system (together with a certain coarse graining) so that it is, in a general sense, assigned to "a description" of the state of the system rather than to the instantaneous real (although perhaps unknown) state of the system and thus has its fundamental significance deeply rooted on an information theoretical setting [8].

Let us note, and warn the reader, that there seems to exits a certain degree of confusion associated with the fact that in a thermodynamical formalism there is in general no explicit procedure to evaluate the entropy of a configuration that is far from equilibrium, a fact that is sometimes misinterpreted as indicating that under those conditions the entropy might not be defined. This is incorrect. Indeed, one of the advantages of statistical mechanics over thermodynamics is that it gives us in principle a procedure to evaluate the entropy of out of equilibrium configurations. Consider, for instance, a gas inside a box, for which the number density and energy density are given at time t in terms of certain functions of the position (corresponding to a situation far from equilibrium). Imagine, moreover, that we are given a certain coarse graining indicating the margin of error incurred in the determination of these functions. Based on the theory of the relevant microscopic degrees of freedom, in this case the kinetic theory of gases, one considers the set of all the microscopic configurations compatible with the data, and, by taking the logarithm of the volume of this set, one can evaluate the corresponding S^{stat}. Note, however, that in this situation we do not expect the result to be a simple function of a few macroscopic parameters, such as the total main energy of the gas and volume of the box.

Let us now consider the specific case of the entropy assignment for a black hole. Recall that the similarity between the laws of black hole mechanics together with the discovery that (stationary) black holes would radiate with a thermal spectrum whose temperature is proportional to the black hole's surface gravity, have lead to the identification of a quarter of the area A of the black hole horizon with its entropy. The issue we need to address is *under what circumstances should we expect this identification to hold?*

To start with, the study of several *gedanken* experiments involving sending ordinary matter endowed with its ordinary entropy into black holes, indicate that the so–called generalized entropy $S^* = S' + A/4$ satisfies the generalized second law in the sense that S^* is a non-decreasing function "of time", despite all attempts to find counterexamples [9]. This together with the first law leave little doubt that the identification should be valid at least in stationary and quasi-stationary situations. How about non-stationary situations? On the one hand we have the so called "Area Theorem" [10], stating that (under certain reasonable conditions) the total area of the event horizon's intersection with a given Cauchy hypersurface, is always greater or equal than the corresponding intersection area for any earlier Cauchy hypersurface. This suggests that the identification should be valid also in non-stationary situations. On the other hand, as mentioned before, we would not normally expect that the entropy of a non-equilibrium configuration should be given by a simple function of a single macroscopic parameter such as $A/4$. However, we must recognize the possibility that, in some respects, black holes might be, at the fundamental level, "much simpler objects" than, say, a gas cloud, and that for such objects these kind of simple relations might remain valid even outside equilibrium. Moreover, if there was an independent quantity (i.e. the purported expression for S^{stat}), different and independent from $A/4$, satisfying, as does $A/4$, the non-decreasing property, we would be in a very strange situation where, in contrast with what happens with ordinary systems, there would be two independent objects indicating the same arrow of time.

It is then clear that things could become very strange unless we maintain that

(iii) *the identification of the horizon area and the black hole entropy should be valid always* (admittedly, there could be correction terms when the curvature reaches the Plank scale, but this issue is not connected with the situation we will be examining).

3 The Schrödinger Black Hole

We consider next a situation which, together with the discussion above, will leads us to remarkable conclusions about the theory of quantum gravity or, alternatively, to retreat from some of the positions (i), (ii), and (iii) already outlined and argued for. The example will involve a certain similarity with the famous "Schödinger Cat", and thus the name "Schrödinger Black Hole" we are giving to it.

Consider a static spherically symmetric thin shell of mass M, in a similarly static, spherically symmetric asymptotically flat space-time. The shell is fitted with a quantal device that at time $t = 0$ (according to an internal clock) will make a random choice between triggering (with probability p) or not (with probability $q = 1 - p$) the collapse of the shell, which would result in the formation of a black hole. More concretely, imagine the shell as made of two thin massless concentric spherical reflecting walls separated by a small distance, with electromagnetic radiation confined between them. The triggering device is connected to a mechanism that makes the internal wall of the shell transparent to radiation when it is activated (this could be done for instance using polarized radiation and Polaroid material for the shells). Next we describe the quantum triggering mechanism: We want to ensure the synchronization on the change in the transparency of the different parts of the internal spherical wall without having to delay the collapse for a time comparable to the light travel time across the shell after the quantum mechanical decision has been made. This can be achieved, for instance, at two opposite points of the shell by using an EPR device as our trigger mechanism: Take a zero spin particle at the center of the shell, which then decays into two photons. Let's fit the internal wall with detectors that will measure the helicity of the photons and give each of them instructions to change the transparency if the photon it detects has positive helicity, but not if it has negative helicity (in this particular case we have $p = q = 1/2$ but the scheme can be easily generalized adjusting the corresponding phases, to achieve any chosen value of $p \in [0, 1]$). In this way, a coordinated collapse of the shell will start at opposite points in the shell, without the need to propagate signals across the shell after the quantum choice is made. One can easily extend this synchronization to the entire shell by means of correlated many-particle states.

The metric outside the thin shell is, of course, the Schwarzschild metric:

$$ds^2 = -\left(1 - 2\frac{GM}{r}\right) dt^2 + \left(1 - 2\frac{GM}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{1}$$

for $r \geq R_{\text{shell}}$, and the metric inside it is the Minkowski metric:

$$ds^2 = -dT^2 + dR^2 + R^2 d\Omega^2 \tag{2}$$

for $R \leq R_{\text{shell}}$. We approximate the shell as infinitely thin. The matching of the exterior coordinates (t, r) with the interior coordinates (T, R) can be deduced from the requirement that the metric induced on the shell from the exterior space-time must coincide with that induced from the interior space-time; while the trajectory of the shell (in the case where it does move) can be deduced from the requirement that it move at the speed of light, see [11].

The motion of the shell be given by specifying the functions $r_{\text{shell}} = R^{(1)}(t)$, in terms of the exterior coordinates, r, t, and $R_{\text{shell}} = R^{(2)}(T)$, in terms of the interior coordinates, R, T.

A simple analysis [5] leads to $r_{\text{shell}} = R_{\text{shell}}$, or in other words, to $R^{(1)}(t) = R^{(2)}(T)$ and

$$\left[\left(1 - 2\frac{M}{R}\right) - \left(1 - 2\frac{M}{R}\right)^{-1} \left(\frac{dR}{dt}\right)^2\right] dt^2 = \left[1 - \left(\frac{dR}{dT}\right)^2\right] dT^2 . \tag{3}$$

We choose the coordinates so that the "quantum measurement" occurs at $T = t = 0$. Then for $t, T < 0$ the shell is static and we have $R = R_0$, where R_0 is the initial radius of the shell. In this case, we obtain from (3)

$$T = \sqrt{1 - 2\frac{M}{R_0}}\, t . \tag{4}$$

If the shell fails to collapse, then Eq.4 remains true for all time. On the other hand, if the shell collapses as a null shell starting at $t = 0$, we can obtain (for T,$t \geq 0$) both T and t as functions of R from the condition that both sides of Eq.3 vanish, i.e. that the induced metric on the shell be degenerate. From the right hand side of this equation we obtain $R(T) = R_0 - T$, and from its left hand side we find

$$R(t) - R_0 + 2M \log\left(\frac{R(t) - 2M}{R_0 - 2M}\right) = -t\,, \tag{5}$$

where we have used the initial condition $R(t = 0) = R_0$.

We next note that, although the collapsing shell will cross the Schwarzschild radius at $t = +\infty$, $T = R_0 - 2M$, in fact, the horizon will be formed earlier than that. Consider a light signal starting at the center of the shell at $T = T_1$ and traveling radially outwards. It will be able to escape to infinity iff it reaches the shell before the collapse has occurred. That is, it must reach the shell while one still has $R_{\text{shell}} > 2M$. The signal travels according to $R = T - T_1$, whence it will meet the shell when $T - T_1 = R(T) = R_0 - T$, that is to say, at $T = (1/2)(R_0 + T_1)$, at which time $R_{\text{shell}} = (1/2)(R_0 - T_1)$. So, the signal will escape iff $T_1 < R_0 - 4M$. If we take, for example, the initial shell radius to be $R_0 = 3M$, then the signal must leave the center with $T_1 < -M$ to be able to escape, and the origin at $T > -M$ is already inside the horizon, if it turns out that the collapse is in fact triggered at $T = 0$.

To summarize, the locus of the horizon at times earlier than $T = 0$ depends on what happens at $T = 0$. If there is no collapse, there is of course no horizon. If the collapse occurs at $T = 0$ then the locus of the horizon at earlier times is given by

$$T - R = R_0 - 4M\,. \tag{6}$$

Consider now a Cauchy hypersurface Σ_{t_c} defined by the condition $t = t_c$ (with $t_c < 0$) outside the shell and by the corresponding condition $T = T_c = (1 - 2M/R_0)^{1/2} t_c$ inside the shell. What is the area A_{t_c} of the intersection of the horizon with Σ_{t_c}? If we choose $t_c > -(4M - R_0)(1 - 2M/R_0)^{-1/2}$ and assume that at $t = 0$ the collapse is in fact triggered, then we will have

$$A_{t_c} = 4\pi R_c^2 = 4\pi \left(4M - R_0 + \sqrt{1 - 2\frac{M}{R_0}} t_c\right)^2 . \tag{7}$$

For example, if we choose $R_0 = 3M$, then we will have a non-vanishing area for $t_c > -\sqrt{3}M$ (if the collapse is triggered) and its value will be

$$A_{t_c} = 4\pi \left(M + \frac{1}{\sqrt{3}} t_c\right)^2 \tag{8}$$

so that for, say, $t_c = -(\sqrt{3}/2)M$ we will have $A_{t_c} = \pi M^2$. Of course, if at $t = 0$ the collapse is not triggered, we will have $A_{t_c} = 0$. Note that by taking R_0

sufficiently close to $2M$ we can have, assuming that it is possible, in principle, to build a shell arbitrarily close to the Schwarzschild radius, a nonzero intersection of Σ_{t_c} with the horizon as early as desired in exterior time t_c; however, we will always have $T_c > -(4M - R_0) > -2M$ when we have such a non-zero intersection. In all these situations, the area A_{t_c} will be bounded by $16\pi M^2$, of course.

Note that we have chosen a foliation of the space-time that seems very natural and which, in particular, is orthogonal to the static Killing field in the region where such field is present. One cannot claim that every conceivable foliation exhibits the same anticipatory behavior as the two we have studied above. However, the classical second law requires the entropy to increase along an *arbitrarily* well-defined foliation, and if such result is to emerge from the quantum gravitational second law it is natural to expect the latter will have to share that feature. Thus the existence of even one foliation along which the horizon forms before the quantum choice is made should be enough to establish the physical significance of the situation been considered here.

I must mention an interesting issue that arose during this meeting, namely whether the setup envisioned here can, in principle, be constructed out of the materials that we find in nature? That is, whether the spherical shells that must contain the electromagnetic radiation can be made sufficiently light so the calculations presented here are not seriously affected. Although this is an interesting point we do not think that it can have any bearing on the discussion at hand, as we can always place a much heavier and solid shell S_{ext} outside our double shell, which will have no effect on the geometry interior to it, and the only thing we need to ensure is that such a shell has not lead to a collapsed situation (the total mass must be bounded by twice it radius). From S_{ext} we can hang now ropes to sustain the interior (which is the part sustaining higher stresses) of our original shell in order to ensure that it can withstand the radiation pressure. Indeed we can think of the ropes being supported at infinity and do without S_{ext}. Ropes subjected to precisely the same type of conditions are considered in examples designed to provide *gedanken* tests of the generalized second law [9].

Finally we note that despite the similarity of the situation considered here with that of the "Schödinger Cat", there is also a big difference between the two situations: according to the basic principles of quantum mechanics the ambiguity in the magnitude of the horizon area (at t_c) in the above example, is completely objective in nature. That is, it is not that "we avoid finding out" whether a horizon exists or not, but that it is objectively impossible for **anyone** to find out, given access only to information available on the given hypersurface.

4 The Problem and the Lessons

Let us consider what entropy should we assign (as dictated by (i) and (ii)) to the instantaneous situation corresponding to the hypersurface Σ_{t_c}? Following (iii) we would like to assign in this case an entropy given by 1/4 the area of the intersection of the event horizon with Σ_{t_c} (plus the entropy of the radiation itself which we regard as negligible and which has in any event the same value for both of the two alternative developments). But this quantity does not have a definite value! This in itself should make us pause and reassess our convictions.

We can continue the analysis of the situation by noting that the fact that with probability p the horizon area is given by Eq.7, and with probability $q = 1 - p$ its value is 0, there is obvious and natural expectation is for the corresponding entropy:

$$S_{t_c} = p \times \frac{A_{t_c}}{4} + q \times 0 + S' \tag{9}$$

where S' is the entropy of associated with the radiation in the shell and other small corrections. We have in fact provided supporting evidence that this should be the entropy associated with the example considered here, including a direct calculation using a sum over histories approach [5], together with the assumption that the standard identification of event horizon and entropy holds in the "normal" circumstances.

Now, let us see what lessons can we extract from this example. Suppose for a moment that we are in the possession of a fully satisfactory theory of quantum gravity that is canonical in the sense of describing every physical situation in terms of operators representing canonical variables in a Hilbert space and representing the possible states of the system at any given time (we are assuming that there would be a relativistic version of "a given time", which could be some generalization of "a given hyprsurface"). These types of theories include both of the most popular candidates for a theory of Quantum Gravity: String Theory and Loop Quantum Gravity. Now, let us imagine how would we proceed with one such theory, to obtain the expression for the entropy associated with Σ_{t_c}. Recall, that as in the case of the inhomogeneous cloud of dust mentioned in Sect. 2, we are supposed to consider the set of microscopic states, presumably described by our QTG, that are compatible (to within a certain margin of precision, presumably provided together with the data) with the macroscopic description of the state of the system. Then, we should evaluate the entropy by taking for instance the logarithm of the number of such microstates. We note, that the macroscopic description of the state of the system in our case is given by the initial data induced on Σ_{t_c} by the space-time described in our example, and by the state of the matter fields describing the radiation, and the other pieces of the setup including the quantum mechanical triggering device. As such the description is as complete as it can be, but is still not enough to determine whether or not there is a black hole horizon intersecting the hypersurface, and moreover such determination is forbidden by the principles of quantum theory and thus, it is also impossible to make such determination in terms of our QTG. Moreover, if we are to recover, in some approximation, the result of (9), the number of such states will depend very strongly on p. However, the value of p is associated, within the data on Σ_{t_c}, only with an energetically insignificant value of a quantum mechanical phase. Needless is to say, that no theory with this strange characteristic has been proposed so far, and that certainly this is not a feature of either String Theory or of Loop Quantum Gravity. Finally, it is worth pointing out that, when taking the limit $p \to 1$, our situation reduces to a standard type of dynamical black hole, because in every ordinary collapse leading to the development of an event horizon, the locus of the latter is influenced to some degree by the matter that will eventually participate in the collapse. It is thus natural to expect that in such limit the evaluation of the number of corresponding states in the Quantum Theory, would reduce to a calculation that is in principle similar to the ones that have been carried out to date, admittedly, within a more restricting set of assumptions.

Thus we have, apparently, been lead to the conclusion that the successful theory of Quantum Gravity must be of a non-canonical type, and in particular that it should be very different from the currently favored candidates. However, we must as always recognize that such conclusion is certainly no stronger than the assumptions we have made. We might want to consider which ones want to do without. As we have argued, **i)** *and* **ii)** would call into question the foundations of the statistical mechanical approach to the derivation of thermodynamics. We might , on the other hand, want to consider giving up on these, but only within the context of black hole physics, but this would mean we would be calling into question the applicability of statistical mechanics to black hole physics and thus we would be taking the foundations out from under the successes in this field of the two approaches to the QTG, which is, after all, what presumably we are trying to preserve with such a drastic steep). We might consider giving up on **iii)**, and take the position that we should rely not on the event horizon, but on some other concept of horizon, which in simple circumstances coincides with the former. This seems to be the position favored by many colleagues and it certainly cannot be dismissed. However if one wants to make progress one must actively look for such an appropriate alternative concept . Lets us review then the existing options and name some of the reasons that they do not provide fully satisfactory choices.

I) The first possibility, is to replace the Event Horizon by the Apparent Horizon. This option has very serious problems, since it is known to be discontinuous for dynamical situations like a collapsing star [12]. Furthermore, it is known that even the Schwarzschild space-time contains Cauchy hypersurfaces with no Apparent Horizons [13].

II) The second alternative, provided by the Isolated Horizons, is particularly interesting for several reasons. First, it has been shown that for *quasi-stationary* processes, the (quasi-local) Horizon Mass satisfies a first law in which the entropy is proportional to the horizon area. Secondly, there exists a calculation of the Statistical Mechanical Entropy that recovers the "standard result" $S = A/4$ for various types of black holes [14]. This formalism is in fact a generalization of the standard stationary scenario to certain physically realistic situations, in which one does not require the exterior region to be in equilibrium. Nevertheless, the whole approach is based on the assumption that the horizon itself is in internal equilibrium. In particular, its area has to be constant, and nothing can "fall into the horizon". In this regard, isolated horizons as presently understood, are not fully satisfactory since the formalism is not defined and does not work in general, dynamical, situations. Moreover, there are situations in which one is faced with the occurrence of several isolated horizons, intersecting a single hypersurface, one within the other, and one must decide to single out the one to which entropy is to be assigned. We can take the view that this should be the outermost horizon, but this seems to be just an *add hoc* choice, unless it is argued that the selection is the natural one associated with the fact that we are specifying the "exterior" observers to be the ones with respect to which entropy is assigned. This view would be natural if we take the position that the assignment of entropy is related to the coarse graining, which is partially specified by pointing out the region from which information is available to the observer. However, this point of view would conflict with the fact that the isolated horizons are not good indicators of such regions, basically because their definition is purely local and thus not fully based on causal relations.

Isolated horizons are well defined for equilibrium situations. If some matter or radiation falls into the horizon, the previously isolated horizon Δ_0 will cease to be isolated, and (one intuitively expects) there will be in the future a new isolated horizon Δ_1, once the radiation has left and the system has reached equilibrium again. One would like to have a definition of horizon that interpolates between these two isolated horizons Δ_0 and Δ_1, such that the physical situation can be described as a generalized horizon that "grows" whenever matter falls in. There is a natural direction for this notion of horizon, and this leads us to the third possibility, namely, Trapping Horizons.

III) In a series of papers, Hayward [15] has been able to show that there exists (at least in the spherically symmetric case) a dynamical (as opposed to quasi-stationary) first law, for a (quasi-local) energy that, however, does not coincide with the Horizon energy of the isolated horizons formalism (in the static limit). There exists also a second law, for the area of the trapping horizon, where a particular foliation of the space-like horizon is chosen. However, we face the problem that, by definition, these horizons can be specified only when the full space-time is available: given a point in space-time, the issue of whether or not it lies on a marginally trapped 2-surface (a key issue to determine whether the point lies or not on the Trapping Horizon), can not in general, be fully ascertained until the whole space-time (where the rest of the 2-surface is to be located) is given. This option is also problematic because the trapping horizons are in general space-like and thus there is no guarantee that a given hypersurface would not intersect the horizon in several disconnected components thus leading to the same problem of in-definition that was mentioned in connection with option II). Moreover, in this case the horizon can even be tangent to the hypersurface which is an extreme version of the previous problem. The fact that all this objections can be raised against this option, has its origin in the fact that the trapping horizon is not a surface defined on the grounds of causality alone. Thus the exiting alternatives to the Event Horizon, do not seem to provide satisfactory options, but is certainly not *apriori* clear that such option can not be constructed.

In conclusion, the issues associated with the assignment of entropy to the Schrödinger Black Hole seems to give a very strong indication that we must give up, either on the QTG of the canonical type, or on the foundations of statistical mechanics, at least as it applies to black holes, or, minimally, that there is something fundamental that we do not understand about these issues.

Acknowledgements

I would like to thank A. Ashtekar, R. Sorkin, R. Wald, A. Corichi, and several of the colleagues attending this meeting for the many interesting discussions about this subject. This work was in part supported by DGAPA-UNAM Grant No. 112401 and by CONACyT grant 32272-E.

References

1. See for example R.M. Wald: The Thermodynamics of Black Holes, *Living Reviews in Relativity* **2001-6**, Preprint gr-qc/9912119; R.M. Wald: *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*, (University of Chicago Press, 1996) and references therein
2. S.W. Hawking: Particle creation by black holes, *Commun. Math. Phys.* **43**, 199 (1975)
3. A. Strominger and C. Vafa: Microscopic origin of the Bekenstein-Hawking entropy, *Phys. Lett. B* **379**, 99 (1996); J. Maldacena and A. Strominger: Statistical entropy of four-dimensional extremal black holes, *Phys. Rev. Lett.* **77** 428-429 (1996); G. Horowitz: Quantum States of Black Holes, in R. Wald (ed.): *Black Holes and Relativistic Stars*, Chicago University Press, 1998
4. A. Ashtekar, J. Baez, A. Corichi, K. Krasnov: Quantum Geometry and Black Hole Entropy, *Phys. Rev. Lett.* **80**, 904 (1998); A. Ashtekar, J. Baez, K. Krasnov: Quantum Geometry of Isolated Horizons and Black Hole Entropy, *Adv. Theor. Math. Phys.* **4**, 1 (2001), `gr-qc/0005126`
5. R. Sorkin and D. Sudarsky: Large Fluctuations in the Horizon Area and what they can tell us about Entropy and Quantum Gravity, *Class. Quantum Grav.* **16**, 3835 (1999)
6. A. Corichi and D. Sudarsky: When is $S = 1/4A$, submitted to *Modern Physics Letters A*
7. Penrose O.: *Foundations of Statistical Mechanics*, (Pergamon Press, Oxford, 1970)
8. C.E. Shanon and W. Weaver: *The Mathematical Theory of Communications* (Univ. of Illinois Press, Urbana, 1949)
9. S. Gao and R.M. Wald: The physical process' version of the first law and the generalized second law for charged and rotating black holes, *Phys. Rev. D* **64**, 084020 (2001), `gr-qc/0106071`
10. P T. Chrusciel, E. Delay, G.J. Galloway, and R. Howard: The area theorem, `gr-qc/0001003`
11. For the classical analysis of collapsing shells see W. Israel: *Nuovo Cim.* **44 b**, 1 (1966); W. Israel: *Phys. Rev.* **153**, 1388 (1967)

12. S.W. Hawking: The Event Horizon, in *Black Holes*, eds. DeWitt and DeWitt (Gordon and Breach, 1973)
13. R. Wald and V. Iyer: Trapped surfaces in the Schwarzschild geometry and cosmic censorship, *Phys. Rev. D* **44**, R 3719 (1991)
14. A. Ashtekar, C. Beetle, S. Fairhurst: Isolated Horizons: a generalization of Black Hole Mechanics. *Class. Quant. Grav.* **16**, L1-L7 (1999); A. Ashtekar, C. Beetle, S. Fairhurst: Mechanics of Isolated Horizons. *Class. Quant. Grav.* **17**, 253 (2000); A. Ashtekar, A. Corichi, K. Krasnov: Isolated Horizons: The Classical Phase Space. *Adv. Math. Theor. Phys.* **3**, 419 (2000), `gr-qc/9905089`
15. S. Hayward: Black Holes: New Horizons. Talk Given at the 9th Marcel Grossman Meeting. Preprint `gr-qc/0008071`; S. Hayward: Quasilocal first law of black-hole dynamics, *Class. Quantum Grav.* **17**, 2153 (2000); S. Hayward: Dynamic black hole entropy, *Phys. Lett. A* **256**, 347 (1999); S. Hayward: General laws of black hole dynamics, *Phys. Rev. D* **49**, 6467 (1994); S. Hayward: Spin coefficient form of the new laws of black hole dynamics, *Class. Quant. Grav.* **11**, 3025 (1994)

Physics with Large Extra Dimensions and Non-Newtonian Gravity at Sub-mm Distances

Ignatios Antoniadis

CERN, Theory Division, CH–1211, Geneva 23, Switzerland**

Abstract. The recent understanding of string theory opens the possibility that the string scale can be as low as a few TeV. The apparent weakness of gravitational interactions can then be accounted by the existence of large internal dimensions, in the submillimeter region. Furthermore, our world must be confined to live on a brane transverse to these large dimensions, with which it interacts only gravitationally. In my lectures, I describe this scenario which gives a new theoretical framework for solving the gauge hierarchy problem and the unification of all interactions. I also discuss its main properties and implications for observations at both future particle colliders, and in non-accelerator gravity experiments. Such effects are for instance the production of Kaluza–Klein resonances, graviton emission in the bulk of extra dimensions, and a radical change of gravitational forces in the submillimeter range.

1 Introduction

Look in front of you. Now on your side. Next on the top. These are the known spatial dimensions of the Universe: there are just three. Have you ever wondered about the origin of this number? Have you ever thought if there are new dimensions that can escape our observation?

In all physical theories, the number of dimensions is a free parameter fixed to three by observation, with one exception: string theory, which predicts the existence of six new spatial dimensions. This is the only known theory today that unifies the two great discoveries of 20th century: quantum mechanics, describing the behavior of elementary particles, and Einstein's General Relativity, describing gravitational phenomena in our Universe.

String theory replaces all elementary point-particles that form matter and its interactions with a single extended object of vanishing width: a tiny string. Thus, every known elementary particle, such as the electron, quark, photon or neutrino, corresponds to a particular vibration mode of the string (see Fig. 1). The diversity of these particles is due to the different properties of the corresponding string vibrations.

Until now, there is no experimental confirmation of string theory. String theorists feel though that the situation may be similar to the one of Einstein's General Relativity before 1919, when its first experimental test has arrived in the occasion of a total eclipse of sun. String physicists believe today in string theory,

** On leave from CPHT (UMR 7644 of CNRS), Ecole Polytechnique, F–91128 Palaiseau.

I. Antoniadis, Physics with Large Extra Dimensions and Non-Newtonian Gravity at Sub-mm Distances, Lect. Notes Phys. **631**, 337–354 (2003)
http://www.springerlink.com/

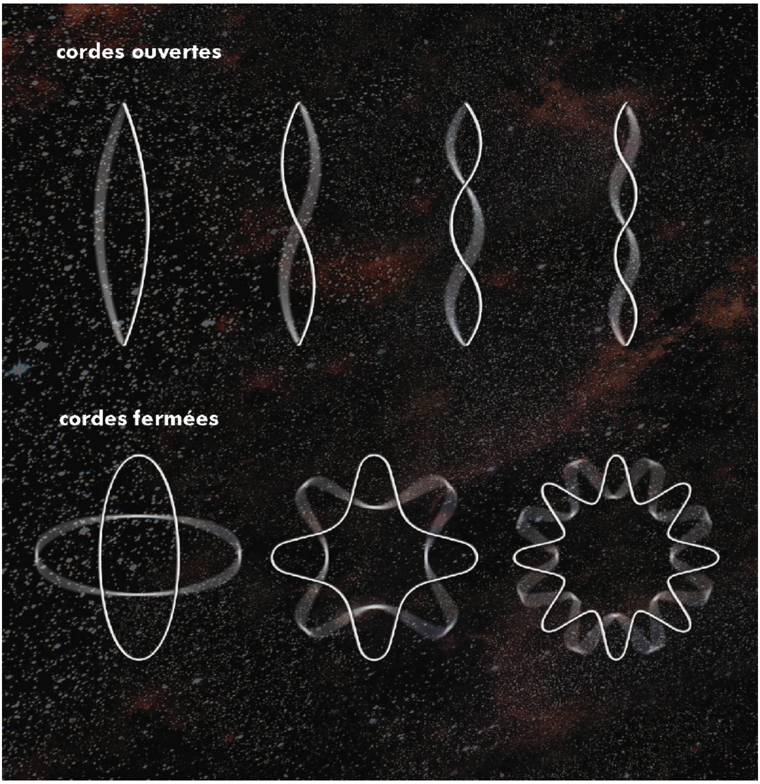


Fig. 1. In string theory, the elementary constituent of matter is a miniscule string, having vanishing width but finite size. It can be open with free ends (*upper part*), or closed (*lower part*). Its vibration modes, like the ones shown above in two dimensions, correspond to various elementary particles

mainly for theoretical reasons, because it provides a framework for unification of all interactions including gravity. However, some precise experimental tests are necessary to decide whether this theory describes the physical reality.

1.1 Small Dimensions

How can it be tested? If our universe has really six additional dimensions, we should observe new phenomena related to the existence of these dimensions. Why nobody has detected them until now? String theorists had an answer for a long time: because the size of the new dimensions is very small, in contrast to the size of the other three that we know, which is infinitely large.

An infinite and narrow cylinder for example is a two-dimensional space, with one dimension forming a very small cycle: one can move infinitely far away along the axis, while one returns back at the same point when moving along the orthogonal direction (see Fig. 2). If one of the three known dimensions of space was small, say of millimeter size, we would be flat and, while we could move

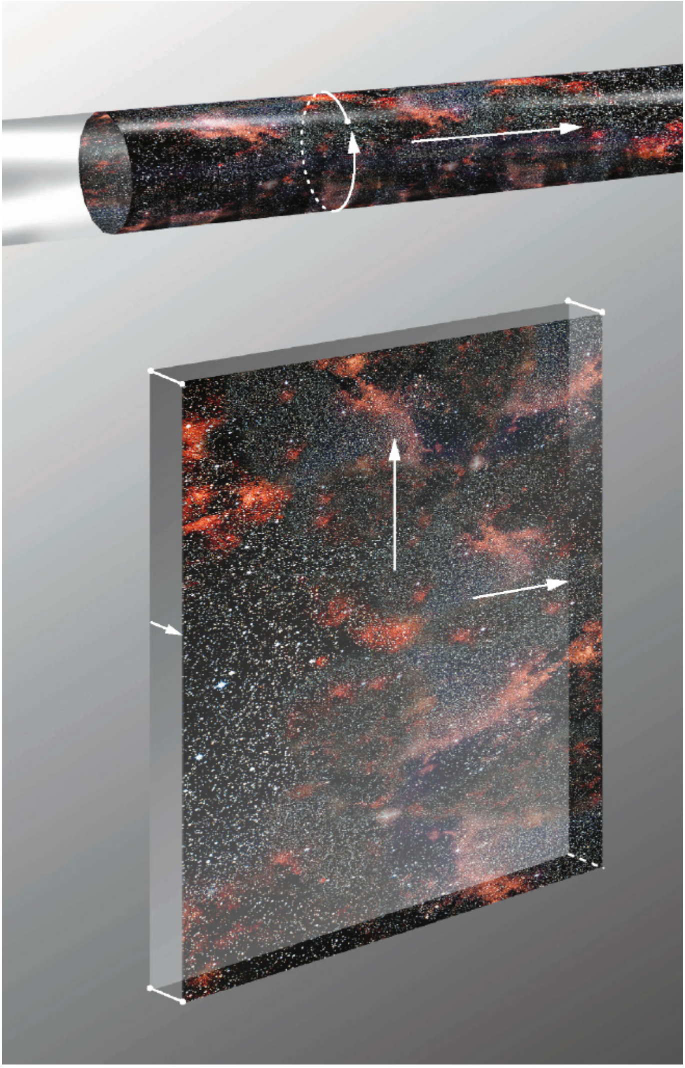


Fig. 2. A small dimension of space may have two different forms. It may close back to itself: in the surface of a cylinder, the dimension transverse to its axis forms a closed cycle. But it may also form an interval, like the thickness of a box: it is impossible to go out of the box, since there is nothing there, not even vacuum. Obviously, it is impossible to draw really extra dimensions within our space

freely towards left or right, forward or backward, it would be impossible to do more than a few millimeters up or down where space ends.

For a long time, string physicists thought that the six extra dimensions were extremely small, having the smallest possible size of physics, associated to the Planck length $\sim 10^{-35}$ meters. In fact, strings were introduced to describe gravitation whose strength becomes important and comparable to the strength of the other three fundamental interactions (electromagnetic, nuclear strong and weak) at very short distances, of the order of the Planck length. It was then natural to assume that the size of the extra dimensions should be of the same order. In this case, the manifestation of new phenomena associated to the extra dimensions are by far out of experimental reach, at least in particle accelerators. Indeed, the Large Hadron Collider (LHC) which is the biggest accelerator under construction at CERN will explore short distances, only up to 10^{-19} meters.

The situation changed drastically recently. During the last three years, more and more theorists examine the possibility that the new dimensions of string theory may be much larger than we thought in the past [1,2]. These ideas lead in particular to experimental tests of string theory that can be performed at TEVATRON and LHC, or at future colliders.

1.2 Supersymmetry Breaking and TeV Dimensions

The first indication of large extra dimensions in string theory came in 1988 from studies of the problem of supersymmetry breaking [3,4]. Supersymmetry is believed to be a new fundamental symmetry of matter which renders the masses of elementary particles compatible with gravitation. In fact, in a quantum theory without supersymmetry, the presence of gravity, which is much weaker than the other interactions, introduces a new high energy scale, the Planck mass $\sim 10^{19}$ GeV, which attracts all masses of elementary particles to become 10^{16} times heavier than their observed values: this is the so-called mass hierarchy problem.

One of the main predictions of supersymmetry is that every known elementary particle has a partner, called superparticle. However, since none of these superparticles have ever been produced in accelerators, they should be heavier than the observed particles. Supersymmetry should therefore be broken. On the other hand, protection of the mass hierarchy requires that its breaking scale, i.e. the mass splitting between the masses of ordinary particles and their partners, cannot be larger than a few TeV. They can therefore be produced for instance at LHC, which will test the idea of supersymmetry.

Assuming that supersymmetry breaking in string theory arises by the process of compactification of the extra dimensions (i.e. from their intrinsic geometry and topology), one can show that its energy breaking scale is tied to the size of these dimensions [3,4]. Thus, a breaking scale in the TeV region would imply the existence of an extra dimension of size $\sim 10^{-18}$ meters. This was one of the few general predictions of string theory which has the chance to be testable in the next generation of particle experiments in the near future.

At that time however, this result was not taken seriously. The reason was rather due to a theoretical prejudice to evade invalidating the only computations we could do. Even now, string theory can be studied in most cases only approximately, namely in perturbation theory. More precisely, computations can be performed if strings interact weakly. The strength of string interactions is controlled by a parameter, called string coupling, which increases when extra dimensions become large. Therefore, when their size is larger than 10^{-35} meters

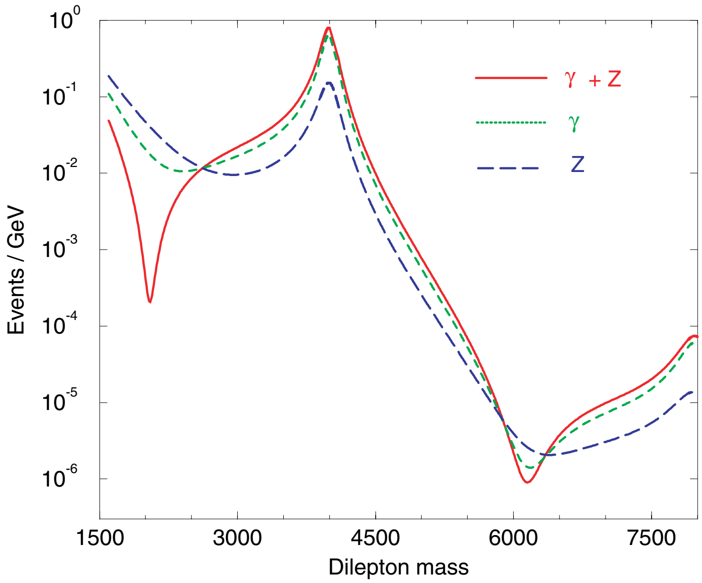


Fig. 3. If there is an extra dimension of size 10^{-18} meters, felt by the electroweak interactions, LHC should produce the first KK states of the photon and of the Z boson. We can then detect the electron-positron pairs produced by the disintegration of these states. The number of the expected events is computed as a function of the energy of the pair in GeV. From highest to lowest: excitation of photon+Z, photon and Z boson

the approximations of perturbative computations do not hold. Thus, the above result has been interpreted negatively. Namely, that supersymmetry breaking could not arise by compactification and remained as an open question.

Two years later specific models were proposed, where perturbative computations were possible in part of the theory, even in the presence of large extra dimensions of size of 10^{-18} meters [1]. This was achieved as a result of conditions imposed to prevent power corrections to low-energy couplings. For gauge couplings, this amounts to the vanishing of the corresponding β-functions, which, for instance, is the case when the so-called Kaluza–Klein (KK) modes are organized in multiplets of $N = 4$ supersymmetry, containing for every massive spin-1 excitation, 2 Dirac fermions and 6 scalars. Examples of such models are provided by orbifolds. In this class of models, the new dimensions form small intervals with our world localized at their ends. The mediators of interactions on the other hand can propagate in the bulk (along the intervals). The study of the physical consequences of these models was performed subsequently in [5–7]. Their main characteristics is the production of the KK states in particle accelerators (see Fig. 3).

If for instance the photon propagates along an extra dimension, one should observe a tower of massive particles with the same properties as the photon but with a mass that becomes larger as the size of the extra dimension is getting smaller. It follows, that for a size of order 10^{-18} meters, an energy of order of a few TeV would be sufficient to produce them. Thus, the existence of such dimensions is testable in LHC. Moreover, these models contain a very light particle that mediates new attractive forces at short distances of the order of a fraction of millimeter, which can be tested in table-top experiments that measure the Newton's law [8] (see Fig. 4). Despite this progress, there was little interest in models with large dimensions because of theoretical reasons related to the large string coupling problem.

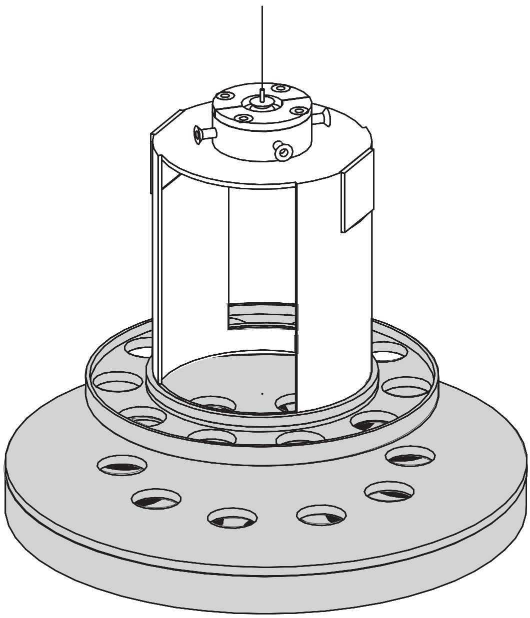


Fig. 4. In his 'Principia' of 1687, Isaac Newton explained that the gravitational attraction between two bodies is proportional to the product of their masses and the inverse square of their distance. The success of this law was spectacular, in particular for the description of the planets motion. At short distances, however, the validity of this law is not tested experimentally. The strongest bounds were obtained using the torsion pendulum shown in the figure, that tested the validity of Newton's law down to 0.2 millimeters. Thus, the size of extra dimensions where gravity propagates is constrained by this value, but allows them to exist at shorter distances, e.g. at 0.01 millimeter. By improving the sensitivity of the measurements, one should then see violations of Newtonian gravity at these distances. The weakness of gravity complicates considerably the experiments: there are several sources of background noise due to other forces that should be eliminated using appropriate devices. At very short distances, one should consider even the Casimir attraction due to fluctuations of the vacuum

1.3 String Dualities

We thus arrived in 1996, when it was realized that the string size $l_s = M_s^{-1}$ is a free parameter of the theory, with a priori no relation to the Planck length [9]. In particular, it could be as large as 10^{-18} meters which is just below the limiting distance that can be probed by present experiments [10]. In order to understand the change of situation, let us return a couple of years earlier. All the works discussed until now were in the context of the so-called heterotic string theory. On the other hand, there were in total five consistent string theories! Four of them contain only closed strings that form closed loops; the other contains also open strings with ends that move freely with the speed of light; besides, all these theories do not have the same amount of supersymmetry.

This multiplicity of theories was creating a problem, since string theory was supposed to provide a unified framework of all physical theories. Thus, one of the five string theories had to be the right one to describe nature, but which one? The answer has been provided in 1994: all of them simultaneously. Following the works of several groups and in particular of Witten, it was realized that every known string theory describes a particular limit of an underlying more general fundamental theory that can be defined in eleven dimensions of spacetime, called M-theory [11].

This discovery made an important progress but did not solved all problems. The main achievement was the connection of the five string theories due to the existence of duality symmetries. One type of these symmetries relates two theories with mutually inverse string couplings. Thus, to solve a problem in the context of some theory with large coupling, it is sufficient to perform an appropriate duality transformation. One obtains then a new problem in the context of a dual theory which has a small coupling, the inverse of the former. The new problem can be solved in perturbation theory of the small coupling. Finally, the resulting solution can be transformed back using the inverse duality transformation that takes us in the first theory. Since computations with large coupling became effectively possible, the road was open to study models with extra dimensions much larger than the Planck length.

1.4 The Universe on a Membrane

A particularly attractive scenario is when the string scale is in the TeV region, which stabilizes the mass hierarchy problem without need of supersymmetry [2]. A possible realization of this idea without experimental conflict is in models possessing large extra dimensions along which only gravity propagates: gravity appears to us very weak at macroscopic scales because its intensity is spread in the “hidden” extra dimensions. On the other hand, at TeV energies, it becomes comparable in strength with the other interactions, i.e. 10^{32} times stronger than what we believed in the past. In order to increase the gravitational force without contradicting present observations, one has to introduce at least two such extra dimensions of size that can be as large as a fraction of a millimeter. At these distances, gravity should start deviate from Newton’s law, which may be possible to explore in laboratory experiments [12] (see Fig. 4).

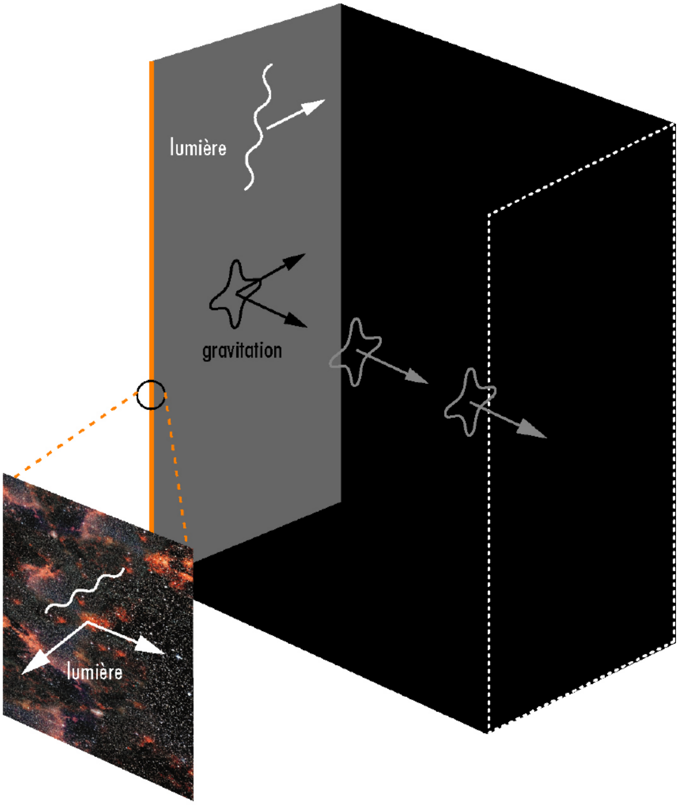


Fig. 5. In the type I string models we consider, our Universe contains, besides the three known spatial dimensions (reduced here to a single orange line), some extra dimensions longitudinal to our world brane (here only one is shown on the grey plane) along which the mediators of gauge interactions described by open strings propagate, as well as some transverse dimensions (here only one again) where only gravity described by closed strings can propagate. Moreover, matter is localized everywhere and propagates only in our three dimensions. The longitudinal extra dimensions have the string size of the order of 10^{-18} meters, while the size of the transverse dimensions varies in the range of 10^{-12} meters and a fraction of a millimeter

A convenient perturbative framework realizing this idea is one of the five string theories, called type I, that contains simultaneously closed and open strings [2,13,14]. Our universe should be localized on a hypersurface, i.e. a membrane extended in p spatial dimensions with $p < 7$, called p-brane (see Fig. 5). Closed strings describe gravity and propagate in all nine dimensions of space: in those extended along the p-brane, as well as in the transverse ones. On the contrary, the endpoints of open strings describing the other (gauge) interactions are confined on the p-brane.

Obviously, our p-braneworld must have at least the three known dimensions of space. But it may contain more: as opposed to the transverse dimensions

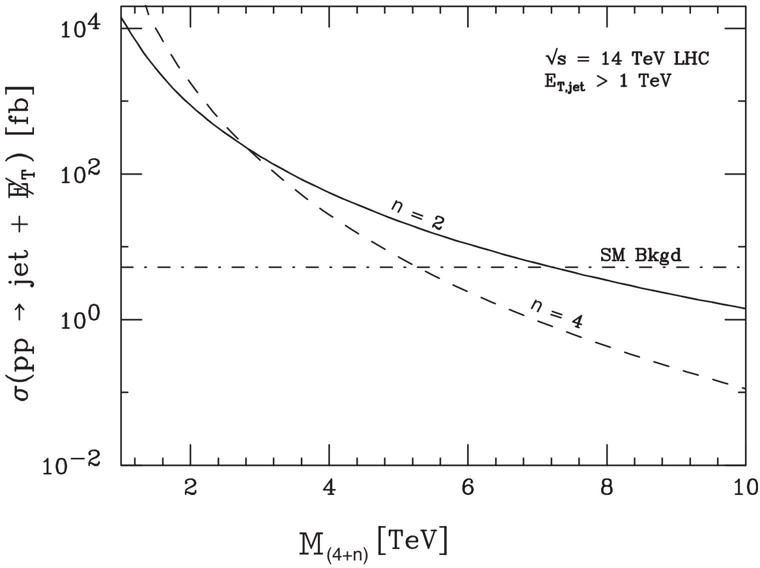


Fig. 6. Missing energy due to graviton emission in the LHC experiment, as a function of the fundamental scale $M_{(4+n)}$ of quantum gravity that propagates in n large transverse dimensions. It is produced together with a hadronic jet that one detects in the collision of the two proton beams. The figure shows the expected cross-section for $n = 2$ and $n = 4$ extra dimensions, together with the background (horizontal dotted-dashed line) coming from other known sources

that interact with us only gravitationally, the "longitudinal" to the brane extra dimensions can be "seen" by the light at sufficiently high energies, giving rise to the production of massive KK particles in accelerators (see Fig. 3).

On the other hand, the existence of the extra large (sub)millimeter dimensions, transverse to our p-brane universe, guarantee that gravitational interactions appear to us very weak at macroscopic distances, larger that a millimeter. The size of these transverse dimensions varies from a fraction of millimeter (in the case of two) to a Fermi (10^{-14} meters, in the case of six). Their characteristic signal in particle colliders is graviton emission into the bulk, leading to missing energy that escapes detection [2,15] (see Fig. 6).

1.5 New versus Old Models

As mentioned before, heterotic string models with large dimensions suffer from a strong coupling problem. When the size of some internal dimensions becomes much larger than the heterotic string length, the string coupling becomes strong invalidating most of the perturbative computations. Now these models may be studied upon an appropriate duality transformation in perturbation theory. Surprisingly, it turns out that most of the old models are in fact equivalent to the recent ones with open strings and transverse dimensions [16]. In addition, by varying the number of large dimensions in the old models, one discovers, due to duality, an alternative way to lower the string scale at the TeV, and thus, accounting for the hierarchy without supersymmetry. In these models, there are no extra large transverse dimensions but still the gravitational force is very weak: it is fixed by the value of the string coupling, which in this case is an independent free parameter of the theory [16,17]. These models have a natural realization in the context of another string theory, called type II.

2 Hiding Extra Dimensions

2.1 Compactification on Tori and Kaluza–Klein States

There is a simple and elegant way to hide the extra dimensions: compactification. It is simple because it relies on an elementary observation. Suppose that the extra dimensions form, at each point of our four-dimensional space, a D-dimensional torus of volume $(2\pi)^D R_1 R_2 \cdots R_D$. The $(4+D)$-dimensional Poincaré invariance is replaced by a four-dimensional one times the symmetry group of the D-dimensional space which contains translations along the D extra directions. The $(4+D)$-dimensional momentum satisfies the mass-shell condition $P^2_{(4+D)} = p_0^2 - p_1^2 - p_2^2 - p_3^2 - \sum_i p_i^2 = m_0^2$ and looks from the four-dimensional point of view as a (squared) mass $M^2_{\mathrm{KK}} = p_0^2 - p_1^2 - p_2^2 - p_3^2 = m_0^2 + \sum_i p_i^2$. Assuming periodicity of the wave functions along each compact direction, one has $p_i = n_i / R_i$ which leads to

$$M^2_{\mathrm{KK}} \equiv M^2_{\boldsymbol{n}} = m_0^2 + \frac{n_1^2}{R_1^2} + \frac{n_2^2}{R_2^2} + \cdots + \frac{n_D^2}{R_D^2}\,, \tag{1}$$

with m_0 being the higher-dimensional mass and n_i non-negative integers. The states with $\sum_i n_i \neq 0$ are called KK states. It is clear that getting aware of the ith extra dimension would require experiments that probe at least an energy of the order of $min(1/R_i)$ with sizable couplings of the KK states to four-dimensional matter.

Let us discuss further some properties of the KK states that will be useful for us below. We parametrize the "internal" D-dimensional box by $y_i \in [-\pi R_i, \pi R_i]$, $i = 1, \cdots, D$ while the four-dimensional Minkowski spacetime is spanned by the coordinates x^μ, $\mu = 0, \cdots, 3$. It is useful to choose for the KK wave functions the basis:

$$\Phi^\alpha_{\boldsymbol{n},\boldsymbol{e}}(x^\mu, y_i) = \Phi^\alpha(x^\mu) \prod_i \left[(1-e_i)\cos\Big(\frac{n_i y_i}{R_i}\Big) + e_i \sin\Big(\frac{n_i y_i}{R_i}\Big) \right]\,, \tag{2}$$

where the vector $\boldsymbol{n} = (n_1, n_2, \cdots, n_D)$ gives the energy of the state following (1) while $\boldsymbol{e} = (e_1, \cdots, e_D)$ with $e_i = 0$ or 1 corresponds to a choice of cosine or sine dependence in the coordinate y_i, respectively. The index α refers to other quantum numbers of Φ.

2.2 Orbifolds and Localized States

The simplest example of the models we will be using are obtained by gauging the Z_2 parity: $y_i \to -y_i \bmod 2\pi R_i$. This leads to compactification on segments of size πR_i. In general, the consistency of this "orbifold" projection implies that the Z_2 space parity should be associated with a Z_2 action on the internal quantum numbers α of Φ. As a result one has the following properties:

- Only states invariant under this Z_2 are kept while the others are projected out. There are two classes of states left in the theory: those for which $\Phi^{(\text{even})}(x^\mu)$ is even under the Z_2 action and $e_i = 0$ and those for which $\Phi^{(\text{odd})}(x^\mu)$ is odd and $e_i = 1$. It is important to notice that the latter are not present as light four-dimensional states i.e. they have $\sum_i n_i \neq 0$ and thus always correspond to higher KK states.
- At the boundaries $y_i = 0, \pi R$ fixed by the Z_2 action, new states $\Phi^{(\text{loc})}(x^\mu)$ have to be included. These "twisted" states are localized at the fixed points. They can not propagate in the extra dimension and thus have no KK excitations.
- The odd bulk states $\Phi^{(\text{odd})}(x^\mu)$ $(e_i = 1)$ have a wave function which vanishes (the $\sin(\frac{n_i y_i}{R_i})$ in (1)) at the boundaries. Their coupling to localized states involves a derivative along y_i. For example three boson interactions of the form $\partial_i \phi^{(\text{odd})})\phi^{(\text{loc})}\phi^{(\text{loc})}$ can be non-vanishing.
- The even states, in contrast, can have non-derivative couplings to localized states. The gauge couplings for instance are given by:

$$g_{\boldsymbol{n}} = \sqrt{2} \sum_{\boldsymbol{n}} \exp\left(- \ln \delta \sum_i \frac{n_i^2 l_s^2}{2R_i^2} \right) g_0 \qquad (3)$$

  where $l_s \equiv M_s^{-1}$ is the string length and $\delta > 1$ is a model dependent number ($\delta = 2^D$ in the case of Z_2's). The $\sqrt{2}$ comes from the relative normalization of the $\cos(\frac{n_i y_i}{R_i})$ wave function with respect to the zero mode while the exponential damping is a result of tree-level string computations.
- Here we are interested in string vacua where gauge degrees of freedom are localized on $(3 + d_\|) + 1$-dimensional subspaces: $(3 + d_\|)$-branes. From the point of view of $(3 + d_\|) + 1$-dimensions the gauge bosons behave as "untwisted" (not localized) particles. In contrast, there are two possible choices for light matter fields. In the first case, they arise from light modes of open strings with both ends on the $(3 + d_\|)$-branes, thus in their interactions they conserve momenta in the $d_\|$ directions. In the second case, they live on the intersection of the $(3 + d_\|)$-branes with some other branes that do not contain the $d_\|$ directions in their worldvolume. These states are localized in the $d_\|$-dimensional space and do not conserve the momenta in these directions. They have no KK excitations and behave as the Z_2 twisted (boundary) states of heterotic strings on orbifolds. The boundary states couple to all KK-modes of gauge fields as described by (3). These couplings violate obviously momentum conservation in the compact directions and make all massive KK excitations unstable.

Use of compactification is an elegant way to hide extra dimensions because some of the quantum numbers and interactions of the elementary particles could be accounted for by the topological and geometrical properties of the internal space. For instance chirality, the number of families in the standard model, gauge and supersymmetry breaking as well as as some selection rules in the interactions of light states could be reproduced through judicious choice of more complicated internal spaces.

3 Low-Scale Strings

In ten dimensions, every superstring theory has two parameters: a mass (or length) scale M_s (or l_s), and a dimensionless string coupling g_s given by the vacuum expectation value (VEV) of the dilaton field $e^{<\phi>} = g_s$, on which we impose the weakly coupled condition $g_s < 1$. Compactification to lower dimensions introduces other parameters describing for instance volumes and shapes of the internal space. The D-dimensional compactification volume V_D will always be chosen to be bigger than unity in string units, $V_D \geq l_s^D$. This choice can always be done by an appropriate T-duality transformation which inverts the compactification radius.

Upon compactification in $D = 4$ dimensions, the above parameters determine the values at the string scale of the four-dimensional (4d) Planck mass (or length) M_P ($l_\mathrm{P} = M_\mathrm{P}^{-1}$) and gauge coupling g_YM that for phenomenological purposes should have the correct strength magnitude. For instance, generically the four-dimensional Planck mass can be expressed as

$$M_\mathrm{P}^2 \simeq \frac{M_s^6 V_6}{g_s^2} M_s^2\,, \tag{4}$$

where V_6 is the six-dimensional internal volume felt by gravitational interactions while the four-dimensional gauge coupling can be written as

$$\frac{1}{g_\mathrm{YM}^2} \simeq \frac{M_s^d V_d}{g_s^q}\,, \tag{5}$$

where V_d is the d-dimensional internal volume felt by gauge interactions.

In the past, weakly coupled heterotic strings were providing the most promising framework for phenomenological applications. In this case, the standard model was considered as descending from the ten-dimensional E_8 gauge symmetry, and we have $V_d = V_6$, $d = 6$ and $q = 2$. Taking the ratio of the two equations, one finds $M_s^2/M_\mathrm{P}^2 \sim g_\mathrm{YM}^2$. Requiring $g_\mathrm{YM} \sim \mathcal{O}(1)$, it was concluded that both the string scale M_s and the compactification scale $R^{-1} \equiv V_6^{-1/6}$ had to lie just below the Planck scale, at energies $\sim 10^{18}$GeV, far out of reach of any near future experiment.

The situation changed during recent years when it was discovered that string theory provides classical solutions (vacua) where gauge degrees of freedom live on subspaces i.e. $d < D = 6$ along with the possibility of $q \neq 2$. For instance,

$(d,q) = (d_{\|}, 1)$ in type I and $(d,q) = (2,0)$ in type II or weakly coupled heterotic strings with small instantons. In these cases, it is an easy exercise to check that both the string and compactification scales can be made arbitrarily low.

The possibility of decreasing the string scale offers new insights on the physics beyond the standard model. For instance, a string scale at energies as low as TeV, would, in addition to the plethora of experimental signatures, provide a solution to the problem of gauge hierarchy alternative to supersymmetry or technicolor. The hierarchy in gauge symmetry versus fundamental (cut-off) scales is then nullified as the two are of the same order [2].

3.1 Type I String Theory and D-Branes

Type I is a ten-dimensional theory of closed and open unoriented strings. Closed strings describe gravity, while gauge interactions are described by open strings whose ends are confined to propagate on p-dimensional sub-spaces defined as Dp-branes. The internal space has 6 compactified dimensions, $p-3 \equiv d_{\|}$ longitudinal and $9-p$ transverse to the Dp-brane.

The gauge and gravitational interactions appear at different orders in string loops perturbation theory, leading to different powers of g_s in the corresponding effective action:

$$S_I = \int d^{10}x \frac{1}{g_s^2 l_s^8} \mathcal{R} + \int d^{p+1}x \frac{1}{g_s l_s^{p-3}} F^2 \,, \tag{6}$$

The $1/g_s$ factor in front of the gauge kinetic terms corresponds to the lowest order open string diagram represented by a disk.

Upon compactification in four dimensions, the Planck length and gauge couplings are given to leading order by

$$\frac{1}{l_{\mathrm{P}}^2} = \frac{V_{\|} V_{\perp}}{g_s^2 l_s^8} \,, \qquad \frac{1}{g_{\mathrm{YM}}^2} = \frac{V_{\|}}{g_s l_s^{p-3}} \,, \tag{7}$$

where $V_{\|}$ ($V_{\perp}$) denotes the compactification volume longitudinal (transverse) to the Dp-brane. From the second relation above, it follows that the requirements of weak coupling $g_{\mathrm{YM}} \sim \mathcal{O}(1)$, $g_s < 1$ imply that the size of the longitudinal space must be of order of the string length ($V_{\|} \sim l_s^{p-3}$), while the transverse volume $V_{\perp}$ remains unrestricted. Using the longitudinal volume in string units $v_{\|} \gtrsim 1$, and assuming an isotropic transverse space of $n = 9-p$ compact dimensions of radius $R_{\perp}$, we can rewrite these relations as:

$$M_{\mathrm{P}}^2 = \frac{1}{g_{\mathrm{YM}}^4 v_{\|}} M_s^{2+n} R_{\perp}^n \,, \qquad g_s = g_{\mathrm{YM}}^2 v_{\|} \,. \tag{8}$$

From the relations (8), it follows that the type I string scale can be chosen hierarchically smaller than the Planck mass at the expense of introducing extra large transverse dimensions that are felt only by the gravitationally interacting light states, while keeping the string coupling weak [2]. The weakness of 4d gravity compared to gauge interactions (ratio M_W/M_P) is then attributed to the largeness of the transverse space $R_\perp/l_s$.

An important property of these models is that gravity becomes $(4+n)$-dimensional with a strength comparable to those of gauge interactions at the string scale. The first relation of (8) can be understood as a consequence of the $(4+n)$-dimensional Gauss law for gravity, with

$$G_N^{(4+n)} = g_{\rm YM}^4 l_s^{2+n} v_\| \tag{9}$$

the Newton's constant in $4+n$ dimensions.

Taking the type I string scale M_s to be at 1 TeV, one finds a size for the transverse dimensions $R_\perp$ varying from 10^8 km, .1 mm (10^{-3} eV), down to .1 Fermi (10 MeV) for $n=1,2$, or 6 large dimensions, respectively. This shows that while $n=1$ is excluded, $n\geq 2$ is allowed by present experimental bounds on gravitational forces [12]. Thus, in these models, gravity appears to us very weak at macroscopic scales because its intensity is spread in the "hidden" extra dimensions. At short distances, gravity should start deviate from Newton's law, which may be possible to explore in laboratory experiments (see Fig. 4).

4 Gravity Modification and Sub-millimeter Forces

Besides the spectacular experimental predictions in particle accelerators, string theories with large volume compactifications and/or low string scale predict also possible modifications of gravitation in the sub-millimeter range, which can be tested in "table-top" experiments that measure gravity at short distances. There are three categories of such predictions:

(i) Deviations from the Newton's law $1/r^2$ behavior to $1/r^{2+n}$, for n extra large transverse dimensions, which can be observable for $n=2$ dimensions of (sub)-millimeter size. This case is particularly attractive on theoretical grounds because of the logarithmic sensitivity of Standard Model couplings on the size of transverse space [14], which allows to determine the desired hierarchy [18], but also for phenomenological reasons since the effects in particle colliders are maximally enhanced [15]. Notice also the coincidence of this scale with the possible value of the cosmological constant in the universe that recent observations seem to support.

(ii) New scalar forces in the sub-millimeter range, motivated by the problem of supersymmetry breaking, and mediated by light scalar fields φ with masses [19,8,2,20]:

$$m_\varphi \simeq \frac{m_{\rm susy}^2}{M_{\rm P}} \simeq 10^{-4} - 10^{-6}\ {\rm eV}\,, \tag{10}$$

for a supersymmetry breaking scale $m_{\rm susy}\simeq 1-10$ TeV. These correspond to Compton wavelengths in the range of 1 mm to 10 μm. $m_{\rm susy}$ can be either the KK scale $1/R$ if supersymmetry is broken by compactification [8,19], or the string scale if it is broken "maximally" on our world-brane [2,20]. A model independent and universal attractive scalar force is mediated by the radius modulus (in Planck units)

$$\varphi \equiv \ln R\,, \tag{11}$$

with R the radius of the longitudinal ($\|$) or transverse ($\perp$) dimension(s), respectively. In the former case, the result (10) follows from the behavior of the vacuum energy density $\Lambda \sim 1/R_{\|}^4$ for large $R_{\|}$ (up to logarithmic corrections). In the latter case, supersymmetry is broken primarily on the brane only, and thus its transmission to the bulk is gravitationally suppressed, leading to masses (10). Note that in the case of two-dimensional bulk, there may be an enhancement factor of the radion mass by $\ln R_{\perp} M_s \simeq 30$ which decreases its wavelength by roughly an order of magnitude [18].
The coupling of the radius modulus (11) to matter relative to gravity can be easily computed and is given by:

$$\sqrt{\alpha_\varphi} = \frac{1}{m}\frac{\partial m}{\partial \varphi} \quad ; \quad \alpha_\varphi = \begin{cases} \dfrac{\partial \ln \Lambda_{\mathrm{QCD}}}{\partial \ln R} \simeq \frac{1}{3} & \text{for } R_{\|} \\ \dfrac{2n}{n+2} = 1 - 1.5 & \text{for } R_{\perp} \end{cases}\,, \tag{12}$$

where m denotes a generic physical mass. In the upper case of a longitudinal radius, the coupling arises dominantly through the radius dependence of the QCD gauge coupling [8], while in the lower case of transverse radius, it can be deduced from the rescaling of the metric which changes the string to the Einstein frame and depends on the dimensionality of the bulk n (varying from $\alpha = 1$ for $n = 2$ to $\alpha = 1.5$ for $n = 6$) [18]. Moreover, in the case of $n = 2$, there may be again model dependent logarithmic corrections of the order of $(g_s/4\pi)\ln RM_s \simeq \mathcal{O}(1)$. Such a force can be tested in microgravity experiments and should be contrasted with the change of Newton's law due to the presence of extra dimensions that is observable only for $n = 2$ [12]. In principle there can be other light moduli which couple with even larger strengths. For example the dilaton, whose VEV determines the (logarithm of the) string coupling constant, if it does not acquire large mass from some dynamical supersymmetric mechanism, can lead to a force of strength 2000 times bigger than gravity [21].

(iii) Non universal repulsive forces much stronger than gravity, mediated by possible abelian gauge fields in the bulk [22,23]. Such gauge fields may acquire tiny masses of the order of M_s^2/M_P, as in (10), due to brane localized anomalies [23]. Although the corresponding gauge coupling is infinitesimally small, $g_A \sim M_s/M_P \simeq 10^{-16}$, it is still bigger than the gravitational coupling $\sim E/M_P$ for typical energies E of the order of the proton mass, and the strength of the new force would be $10^6 - 10^8$ stronger than gravity. This is an interesting region which will be soon explored in micro-gravity experiments (see Fig. 7). Note that in this case the supernova constraints impose that there should be at least four large extra dimensions in the bulk [22].

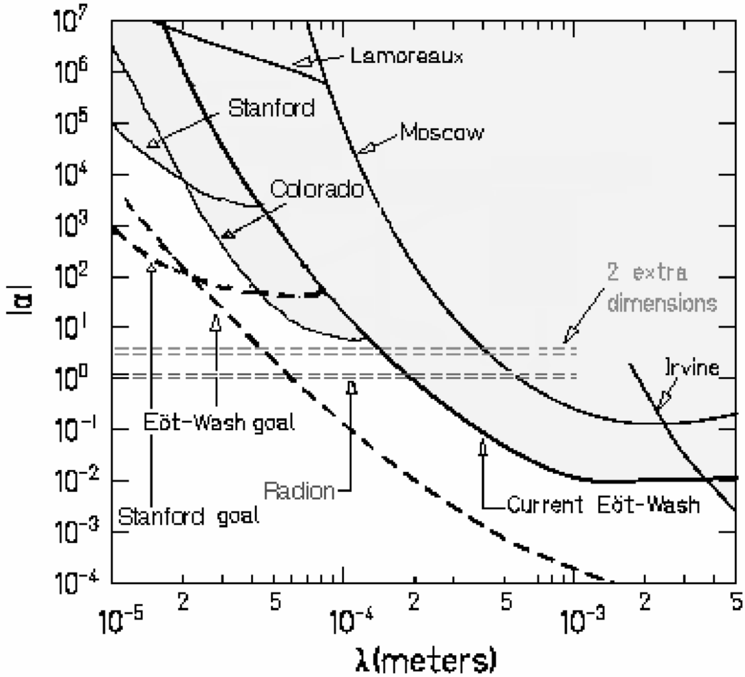


Fig. 7. Limits on non-Newtonian forces at short distances, compared to new forces mediated by the graviton in the case of two large extra dimensions, and by the radion

In Fig. 7 we depict the actual information from previous, present and upcoming experiments [18]. The vertical axis is the strength, α, of the force relative to gravity; the horizontal axis is the Compton wavelength, λ, of the exchanged particle. The solid lines indicate the present limits from the experiments indicated. The excluded regions lie above these solid lines. Measuring gravitational strength forces at such short distances is quite challenging. The most important background is the van der Waals force which becomes equal to the gravitational force between two atoms when they are about 100 microns apart. Since the van der Waals force falls off as the 7th power of the distance, it rapidly becomes negligible compared to gravity at distances exceeding 100 μm. The dashed thick lines give the expected sensitivity of the present and upcoming experiments, which will improve the actual limits by roughly two orders of magnitude, while the horizontal dashed lines correspond to the theoretical predictions for the graviton in the case of two large extra dimensions and for the radion in the case of transverse radius.

5 Conclusions

Clearly, today, these theories exist only in our imagination. However, we look forward at the next generation of high energy experiments and in particular at the most powerful machine, the LHC at CERN. I am convinced, as the majority of my colleagues, that LHC will play a very important role for the future of high-energy physics of fundamental interactions. In fact, it is designed since last decade to explore the origin of mass of elementary particles and to test, in particular, the idea of supersymmetry, looking for the production of superparticles. We now hope that this accelerator may discover more spectacular and “exotic” phenomena, such as the existence of large extra dimensions of space and of fundamental strings.

Acknowledgments

This work was partly supported by the European Commission under RTN contract HPRN-CT-2000-00148 and by the INTAS contract N 99-1-590.

References

1. I. Antoniadis: *Phys. Lett. B* **246**, 377 (1990)
2. N. Arkani-Hamed, S. Dimopoulos and G. Dvali: *Phys. Lett. B* **429**, 263 (1998); I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali: *Phys. Lett. B* **436**, 263 (1998)
3. I. Antoniadis, C. Bachas, D. Lewellen and T. Tomaras: *Phys. Lett. B* **207**, 441 (1988)
4. C. Kounnas and M. Porrati: *Nucl. Phys. B* **310**, 355 (1988); S. Ferrara, C. Kounnas, M. Porrati and F. Zwirner, *Nucl. Phys. B* **318**, 75 (1989)
5. I. Antoniadis and K. Benakli: *Phys. Lett. B* **326**, 69 (1994)
6. I. Antoniadis, K. Benakli and M. Quirós: *Phys. Lett. B* **331**, 313 (1994) and *Phys. Lett. B* **460**, 176 (1999); P. Nath, Y. Yamada and M. Yamaguchi: *Phys. Lett. B* **466**, 100 (1999); T.G. Rizzo and J.D. Wells: *Phys. Rev. D* **61**, 016007 (2000); T.G. Rizzo: hep-ph/9909232; A. De Rujula, A. Donini, M.B. Gavela and S. Rigolin: *Phys. Lett. B* **482**, 195 (2000)
7. E. Accomando, I. Antoniadis and K. Benakli: *Nucl. Phys. B* **579**, 3 (2000)
8. I. Antoniadis, S. Dimopoulos and G. Dvali: *Nucl. Phys. B* **516**, 70 (1998)
9. E. Witten: *Nucl. Phys. B* **471**, 135 (1996)
10. J.D. Lykken: *Phys. Rev. D* **54**, 3693 (1996)
11. E. Witten: *Nucl. Phys. B* **443**, 85 (1995)
12. C.D. Hoyle, U. Schmidt, B.R. Heckel, E.G. Adelberger, J.H. Gundlach, D.J. Kapner and H.E. Swanson: *Phys. Rev. Lett.* **86**, 1418 (2001); J. Chiaverini, S. J. Smullin, A.A. Geraci, D.M. Weld and A. Kapitulnik: hep-ph/0209325; J.C. Long, H.W. Chan, A.B. Churnside, E.A. Gulbis, M.C. Varney and J C. Price: hep-ph/0210004; D.E. Krause, E. Fischbach: in C. Lämmerzahl, C.W.F. Everitt, F.W: Hehl (eds.), *Gyros, Clocks, Interfermeters,...: Testing Relativistic Gravity in Space* (Lect. Notes Phys. **562**), Berlin 2001, p. 292, [hep-ph/9912276]; H. Abele, S. Haeßler and A. Westphal: *these proceedings*
13. G. Shiu and S.-H.H. Tye: *Phys. Rev. D* **58**, 106007 (1998); Z. Kakushadze and S.-H.H. Tye: *Nucl. Phys. B* **548**, 180 (1999); L.E. Ibáñez, C. Muñoz and S. Rigolin: *Nucl. Phys. B* **553**, 43 (1999)
14. I. Antoniadis, C. Bachas: *Phys. Lett. B* **450**, 83 (1999)

15. G.F. Giudice, R. Rattazzi and J.D. Wells: *Nucl. Phys. B* **544**, 3 (1999); E.A. Mirabelli, M. Perelstein and M.E. Peskin: *Phys. Rev. Lett.* **82**, 2236 (1999); T. Han, J. D. Lykken and R. Zhang: *Phys. Rev. D* **59**, 105006 (1999); K. Cheung and W.-Y. Keung: *Phys. Rev. D* **60**, 112003 (1999); C. Balázs et al.: *Phys. Rev. Lett.* **83**, 2112 (1999); L3 Collaboration (M. Acciarri et al.): *Phys. Lett. B* **464**, 135 (1999), *Phys. Lett. B* **470**, 281 (1999); J.L. Hewett: *Phys. Rev. Lett.* **82**, 4765 (1999); D. Atwood, C.P. Burgess, E. Filotas, F. Leblond, D. London and I. Maksymyk: hep-ph/0007178
16. I. Antoniadis and B. Pioline: *Nucl. Phys. B* **550**, 41 (1999)
17. K. Benakli and Y. Oz: *Phys. Lett. B* **472**, 83 (2000); I. Antoniadis, S. Dimopoulos and A. Giveon: *JHEP* **0105**, 055 (2001)
18. I. Antoniadis, K. Benakli, A. Laugier and T. Maillard: CERN-TH/2002-341 preprint to appear
19. S. Ferrara, C. Kounnas and F. Zwirner: *Nucl. Phys. B* **429**, 589 (1994)
20. I. Antoniadis, E. Dudas and A. Sagnotti: *Phys. Lett. B* **464**, 38 (1999); G. Aldazabal and A.M. Uranga: *JHEP* **9910**, 024 (1999); C. Angelantonj, I. Antoniadis, G. D'Appollonio, E. Dudas and A. Sagnotti: *Nucl. Phys. B* **572**, 36 (2000)
21. T.R. Taylor and G. Veneziano: *Phys. Lett. B* **213**, 450 (1988)
22. N. Arkani-Hamed, S. Dimopoulos and G. R. Dvali: *Phys. Rev. D* **59**, 086004 (1999)
23. I. Antoniadis, E. Kiritsis and J. Rizos: *Nucl. Phys. B* **637**, 92 (2002)

Quantum States of Neutrons in the Gravitational Field and Limits for Non-Newtonian Interaction in the Range between 1 µm and 10 µm

Hartmut Abele[1], Stefan Baeßler[2], and Alexander Westphal[3]

[1] Physikalisches Institut der Universität Heidelberg, Philosophenweg 12, 69120 Heidelberg, Germany
[2] Physikalisches Institut der Universität Mainz, Staudingerweg 7, 55099 Mainz, Germany
[3] DESY Hamburg, Notkestraße 85, 22607 Hamburg, Germany

Abstract. Quantum states in the Earth's gravitational field can be observed, when ultra-cold neutrons fall under gravity. In an experiment at the Institut Laue-Langevin in Grenoble, neutrons are reflected and trapped in a gravitational cavity above a horizontal mirror. The population of the ground state and the lowest states follows, step by step, the quantum mechanical prediction. An efficient neutron absorber removes the higher, unwanted states. The quantum states probe Newtonian gravity on the micrometer scale and we place limits for gravity-like forces in the range between 1 µm and 10 µm.

1 A Quantum System

Quantum physics is a fascinating but subtle subject. The subtlety of the quantum system described here arises from the fact, that gravity appears very weak in our universe. Quantum theory and gravitation affect each other, and, when neutrons become ultra-cold, the fall experiment of Galileo Galilei shows quantum aspects of the subtle gravity force in that sense that neutrons do not fall continuously. We find them on different levels, when they come close to a reflecting mirror for neutrons [1]. Of course, such bound states with discrete energy levels are expected when the graviational potential is larger than the energy of the particle. Here, the quantum states have pico-eV energy, a value that is smaller by many orders of magnitude compared with an electromagnetically bound electron in a hydrogen atom, opening the way to a new technique for gravity experiments, for neutron optics, neutron detection and measurements of fundamental properties.

A side-effect of this experiment is its sensitivity for gravity-like forces at length scales below 10 µm. In light of recent theoretical developments in higher dimensional field theory [2–4], gauge fields can mediate forces that are 10^{10} to 10^{12} times stronger than gravity at submillimeter distances, exactly in the interesting range of this experiment and might give a signal in an improved setup.

The idea of observing quantum effects occuring when ultracold neutron are stored on a plane was discussed long ago by V.I. Lushikov and I.A. Frank [5].

H. Abele, S. Baeßler, and A. Westphal, Quantum States of Neutrons in the Gravitational Field and Limits for Non-Newtonian Interaction, Lect. Notes Phys. **631**, 355–366 (2003)
http://www.springerlink.com/

An in some aspects similar experiment was discussed by H. Wallis et al. [6] in the context of trapping atoms in a gravitational cavity. Retroreflectors for atoms have used the electric dipole force in an evanescent light wave [7,8] or they are based on the gradient of the magnetic dipole interaction, which has the advantage of not requiring a laser [9]. A neutron mirror makes use of the strong interaction between nuclei and an ultracold neutron, resulting in an effective repulsive force: Neutrons propagate in condensed matter in a matter similar to the propagation of light but with a neutron refractive index less than unity. Thus, one considers the surface of matter as constituting a potential step of height V. Neutrons with transversal energy $E_\perp < V$ will be totally reflected. Ultra-cold neutrons (UCN) are neutrons that, in contrast to faster neutrons, are retro-reflected from surfaces at all angles of incidence. When the surface roughness of the mirror is small enough, the UCN reflection is specular.

Neutron mirrors are interesting because they can be used to store neutrons, to focus neutrons, or even to build a Fabry Perot interferometer for neutron de Broglie waves. UCN storage bottle experiments have improved our knowledge about the neutron lifetime significantly or, together with the Ramsey method of separated oscillating fields, they have been used for a search for an electric dipole moment of the neutron.

2 Limits for Non-Newtonian Interaction Below 10 µm

Discussions about deviations from the inverse square law for gravity have become popular in the past few years [10]. A new effective interaction coexisting with gravity would modify the Newtonian potential. On the assumption that the form of the non-Newtonian potential is given by the Yukawa expression, for masses m_i and m_j and distance r the modified Newtonian potential $V(r)$ is having the form

$$V(r) = -G\frac{m_i \cdot m_j}{r}\left(1 + \alpha \cdot e^{-r/\lambda}\right) , \tag{1}$$

where λ is the Yukawa distance over which the corresponding force acts and α is a strength factor in comparison with Newtonian gravity. G is the gravitational constant. For a neutron with mass m_n, the gravitational force of the mass m_E of the entire Earth with radius R_E lead to a free fall acceleration

$$g = \frac{Gm_E}{R_E^2} = \frac{4\pi \cdot G\rho R_\mathrm{E}^3}{3R_\mathrm{E}^2} = \frac{4}{3}\pi G\rho R_\mathrm{E} \,. \tag{2}$$

When a neutron approaches the mirror, the mass of this extended source might modify g, when strong non-Newtonian forces are present. For small neutron distances z from the mirror, say several micrometers, we consider the mirror as an infinite half-space with mass density ρ. By replacing the source mass m_i by dm_i and integrating over dm_i, the Yukawa-like term λ modified Newtonian potential $V'(r)$ is having the form

$$V'(z,\lambda) = 2\pi \cdot m_n\rho\alpha\lambda^2 G \cdot e^{-|z|/\lambda} \tag{3}$$

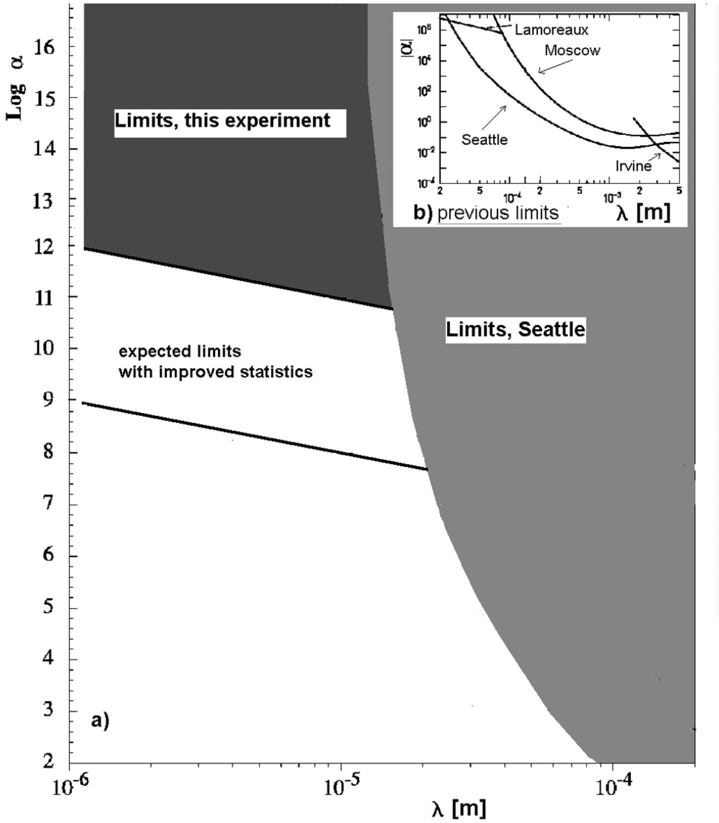


Fig. 1. Limits for non-Newtonian gravity: Strength $|\alpha|$ vs. Yukawa length scale λ. **a** Experiments with neutrons place limits for $|\alpha|$ in the range 1 µm $< \lambda <$ 10 µm. **b** Constraints from previous experiments [11–15] are adapted from [14]

and the non-Newtonian acceleration g' as a function of distance z is given by

$$g'(z,\lambda) = 2\pi \cdot \rho\alpha\lambda G \cdot e^{-|z|/\lambda} \,. \tag{4}$$

As a consequence, the ratio is

$$\frac{g'}{g}(z,\lambda) = \frac{3}{2}\alpha \cdot \frac{\lambda}{R_{\mathrm{E}}} \cdot e^{-|z|/\lambda} . \tag{5}$$

For z $= \lambda = 10\ \mu$m and $\alpha = 1$, the ratio g'/g is about 10^{-12}. Figure 1a shows the present status of an experimental search for gravity-like forces below 10 µm. The results of a fit of potential equation (3) to the measured data (see Figs. 5 and 6) yields predictions for 90% confidence level exclusion bounds on α and λ. These limits from this neutron mirror experiment are shown in Fig. 1a. They are the best known in the range 1 µm $< \lambda <$ 3 µm and exclude for the first time gravity-like short-ranged forces at 1 µm with strength $\alpha > 10^{12}$. The limit for strength α at 10 µm is 10^{11} (Fig. 1a)). For the theoretical biased uncertainty due to the neutron absorber (see Sect. 3.2), we give an estimate of one order of magnitude. In future experiments, these limits will be strongly improved by an enhanced setup and improved statistics by new UCN sources as a Monte Carlo simulation shows. Previous constraints [11–15] on both λ and α are shown in Fig. 1b.

The method with neutrons has some advantages. Electromagnetic interactions at micrometer distances, serious sources of systematic error in distance force measurements, are effectively suppressed. The neutron carries no electric charge and direct electrostatic forces are ruled out. Yet, it bears a very tiny magnetic moment μ_n of roughly $0.5 \cdot 10^{-3} \cdot \mu_B$. This magnetic moment can induce a magnetic mirror force onto a neutron that hovers close to the surface of a body. Further, a neutron moving with the velocity v sees an induced electrostatic force, that is essentially some kind of Lorentz force and thus an effect of the order of v/c. Both effects can be evaluated to yield electrodynamic energy shifts. With permittivity ϵ and permeability μ, the order of magnitude is

$$\Delta E_{\boldsymbol{E},v} \sim \epsilon_0 \cdot \frac{\epsilon - 1}{\epsilon} \cdot \frac{\mu_0^2 \cdot \mu_n^2}{48\pi} \cdot \frac{v^2}{z^3} \sim 10^{-26} \cdot 10^{-12}\,\mathrm{eV} \tag{6}$$

$$\Delta E_{\boldsymbol{B}} \sim \frac{\mu - 1}{\mu} \cdot \frac{\mu_0}{16\pi} \cdot \frac{\mu_n^2}{z^3} \sim 10^{-13} \cdot 10^{-12}\,\mathrm{eV} \tag{7}$$

for $v \sim 10$ m/s and $z \sim 10\ \mu$m. Thus, these effects can completely be neglected, since they are by far subgravitational in strength.

Motivations for gravity experiments come from frameworks for solving the hierarchy problem in a way of bringing quantum gravity down to the TeV scale. In such frameworks the Standard Model fields are localized on a 3-brane in a higher dimensional space by the presence of new dimensions of submillimeter size [2]. At the expected sensitivity, a number of gravity-like phenomena emerge. For example, a hypothetical gauge field can naturally have miniscule gauge coupling $g_4 \sim 10^{-16}$ for 1 TeV, independent of the number of extra dimensions [3]. If these gauge fields couple to a neutron with mass m_n, the ratio of the repulsive force mediated by this gauge field to the gravitational attraction is [3]

$$\frac{F_{\mathrm{gauge}}}{F_{\mathrm{grav}}} \sim \frac{g_4^2}{G m_n^2} \sim 10^6 \left(\frac{g_4}{10^{-16}}\right)^2 . \tag{8}$$

With $g_4 = 10^{-16}$ as a lower bound, these gauge fields can result in repulsive forces of million or billion times stronger than gravity at micrometer distances, exactly in the range of interest.

3 The Experiment at the Institut Laue-Langevin

3.1 From Hot to Ultracold

Neutrons are produced in a spallation source or a research reactor. At production, these neutrons are very hot; the energy is about 2 MeV corresponding to

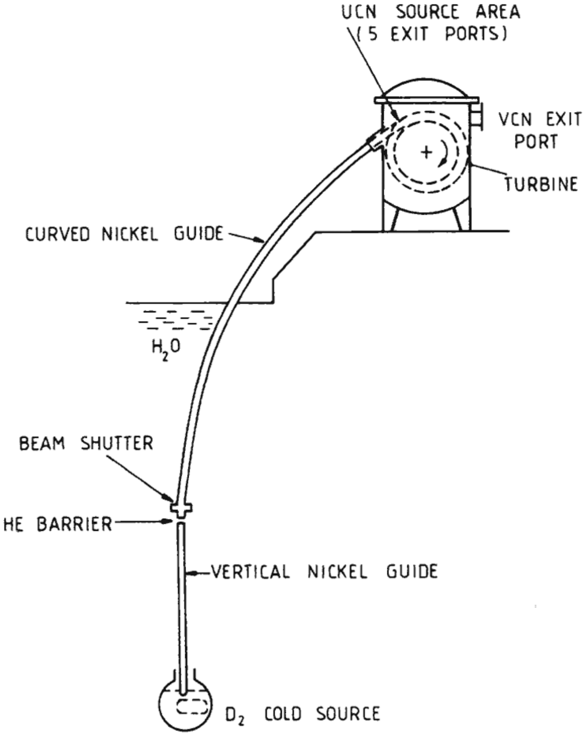


Fig. 2. ucn source

10^{10} degrees Centigrade. On the other side of the scale, the gravity experiment uses neutrons having 10^{18} times less energy in the pico-eV range (see Table 1). In a first step, spallation or fission neutrons thermalize in a heavy water tank at a temperature of 300 K. The thermal fluxes are distributed in energy according to Maxwellian law. At the Institut Laue-Langevin (ILL), cold neutrons are obtained in a second moderator stage from a 25 K liquid deuterium cold moderator near the core of the 57 MW uranium reactor. These cold neutrons have a velocity spectrum in the milli-eV energy range. For particle physics, a new beam line with a flux of more than 10^{10} cm^{-2}s^{-1} over a cross section of 6 cm $\times$ 20 cm is available.

Ultra-cold neutrons are taken from the low energy tail of the cold Maxwellian spectrum. They are guided vertically upwards by a neutron guide (Fig. 2). The curved guide, which absorbs neutrons above a threshold energy, acts as a low-velocity filter for neutrons. Neutrons with a velocity of up to 50 m/s arrive at a rotating nickel turbine. Colliding with the moving blades of the turbine, ultra-cold neutrons exit the turbine with a velocity of several meters per second. They are then guided to several experimental areas. The exit window of the guide for the gravity experiment has a rectangular shape with the dimensions of 100 mm $\times$ 10 mm. At the entrance of the experiment, a collimator absorber system cuts

Table 1. From hot to ultracold: neutrons at the ILL

	fission neutrons	thermal neutrons	cold neutrons	ultracold neutrons	this experiment
Energy	2 MeV	25 meV	3 meV	100 neV	1.4 peV
Temperature	10^{10} K	300 K	40 K	1 mK	-
Velocity	10^{7} m/s	2200 m/s	800 m/s	5 m/s	$v_{\perp} \sim 2$ cm/s

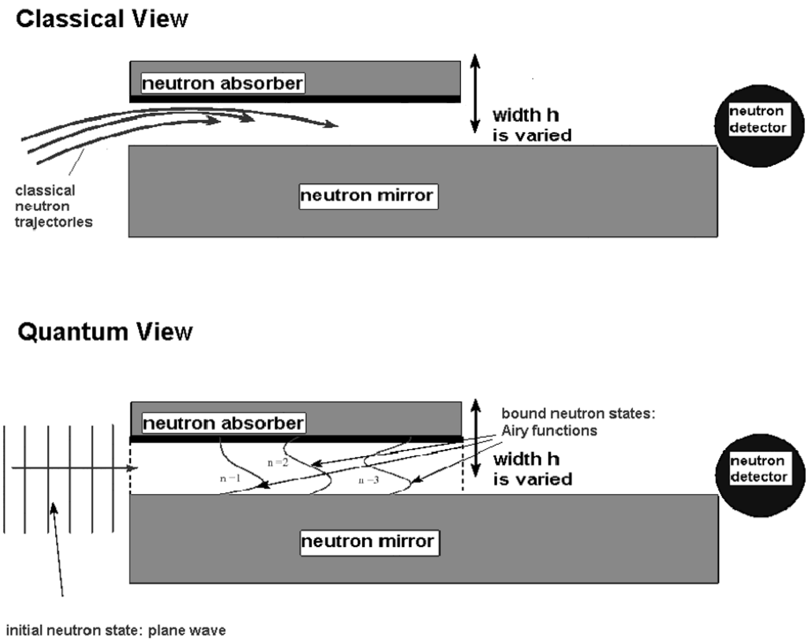


Fig. 3. Sketch of the setup: **a** classical view: neutron trajectories, **b** quantum view: plane waves and Airy functions

down on the neutrons to a adjustable transversal velocity corresponding to an energy in the pico-eV range.

3.2 The Setup

Figure 3 shows a schematic view of the setup: Neutrons pass through the mirror absorber system eventually detected by a ^{3}He-counter. The experiment itself is mounted on a polished plane granite stone with a passive antivibration table underneath. This stone is leveled with piezo translators. Inclinometers together with the piezo translators in a closed loop circuit guarantee leveling with an absolute precision better than 10 µrad. Either one solid block with dimensions 10 cm $\times$ 10 cm $\times$ 3 cm or two solid blocks with dimensions 10 cm $\times$ 6 cm $\times$ 3 cm

Table 2. Eigenenergies and classical turning points for neutrons, atoms and electrons, a comparison

	Neutron	^{4}Helium	^{85}Rubidium	^{133}Cesium	Electron
E_1 $[peV]$	1.4	2.3	6.2	7.2	0.12
E_2 $[peV]$	2.5	3.9	11.0	12.7	0.20
z_1 $[\mu]$	13.7	5.5	0.7	0.5	2061
z_2 $[\mu]$	24.0	9.5	1.2	0.9	3604

composed of optical glass serve as mirrors for UCN neutron reflection. Small angle X-ray studies [16] determined the roughness of the surface to be $\sigma = 2.2 \pm 0.2$ nm and the associated lateral correlation length to be $\zeta = 10 \pm 2$ µm. The plane can thus be regarded as a pattern that varies in height with 2 nm on a scale of 10 µm. Since the de Broglie wavelength of the neutrons is in the range of 40 nm to 100 nm, the neutrons do see a surface that is essentially flat. A neutron-absorber is placed above the first mirror. The absorber consists of a rough glass plate coated with an Gd-Ti-Zr alloy by means of magnetron evaporation. The absorbing layer is 200 nm thick. The surface of the absorber was parallel to the surface of the mirror. The absorber roughness and correlation length was measured with an atomic force microscope to be $\sigma = 0.75$ µm and $\zeta = 5$ µm respectively. Neutrons that hit the absorber surface are either absorbed in the alloy or scattered out of the experiment at large angles. The efficiency of removing these fast unwanted neutrons is 93%. The collimation system in front of the mirror absorber system is adjusted in that way, that classical trajectories of neutrons entering the experiment have to hit the mirror surface at least two times. After the second mirror we placed a ^{3}He counter for neutron detection. More information about the setup can be found in [17].

4 Gravity and Quantum Mechanics Work Together

4.1 Theoretical Description

The neutrons fall under gravity onto the mirror. The calculation of the energy eigenvalues of the vertical motion of the neutrons in the mirror-absorber system is a nice example of quantum mechanics. In fact, we have two theoretical descriptions for the transmission of neutrons. The first one is the well known WKB method. Usually, the accuracy of WKB quantization is 20% for the ground state, whereas the accuracy increases for higher levels. A similar calculation of energy levels for the gravitational field with the WKB method can also be found in [6]. We can compare the WKB result with an exact analytical solution using Airy-functions. Taking the neutron-absorber into account, the agreement of the two methods is significantly better than 10%. We start calculations from the one dimensional stationary Schrödinger equation,

$$-\frac{\hbar^2}{2m} \triangle \Psi + V(z) = E\Psi \tag{9}$$

with wave function Ψ for energy E and the potential

$$V(z) = \begin{cases} mgz & \text{for } z \geq 0\,, \\ \infty & \text{for } z < 0\,, \end{cases} \tag{10}$$

ignoring the absorber for now. m is the mass of the neutron and g is the acceleration in the earth's gravitational field. The quantum mechanical treatment of reflecting neutron mirror, made from glass, is simple. The glass potential is essential real because of the small absorption cross section of glass and with $V = 100\ neV$ large compared with transversal energy $E_\perp$. Therefore, the potential V is set to infinity at z = 0. The quantum mechanical description follows in part [6]. It is convenient to use a scaling factor

$$\zeta = \frac{z}{z_0} \quad \text{with} \quad z_0 = \left(\frac{\hbar^2}{2m^2 g}\right)^{1/3} . \tag{11}$$

Solutions of (9) for Ψ are obtained with an Airy function

$$\Psi_n(\zeta) = \mathrm{Ai}(\zeta - \zeta_n)\,. \tag{12}$$

The displacement ζ_n of the n-th eigenvector has to coincide with the n-th zero of the Airy function (Ai(-ζ_n)=0) to fulfill the boundary condition $\Psi_n(0) = 0$ at the mirror. Eigenfunctions ($n > 0$) are

$$\mathrm{Ai}(\zeta - \zeta_n) \tag{13}$$

with corresponding eigenenergies

$$E_n = mgz_n \tag{14}$$

and

$$z_n = z_0 \left(\frac{3\pi}{2}\left(n - \tfrac{1}{4}\right)\right)^{2/3} . \tag{15}$$

z_n corresponds to the turning point of a classical neutron trajectory with energy E_n. For example, Energies of the lowest levels (n = 1, 2, 3, 4) are 1.44 peV, 2.53 peV 3.42 peV and 4.21 peV. The corresponding classical turning points z_n are 13.7 µm, 24.1 µm, 32.5 µm and 40.1 µm (see Table 2).

$$\rho = C\Psi^*\Psi \tag{16}$$

is the neutron density and can be interpreted as the probability to detect a neutron at height z above the mirror, see Fig. 4. C is a constant.

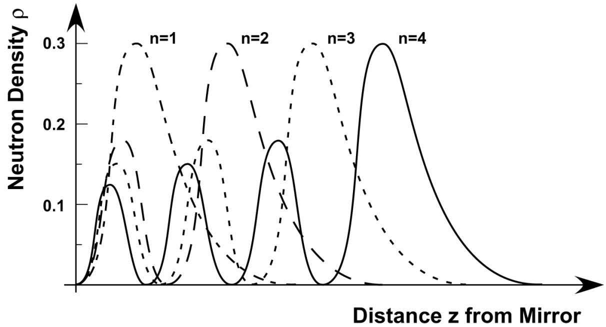


Fig. 4. Neutron density above the mirror for states #1 to #4

In principle it is possible to visualize this neutron density distribution. The distribution is measurable with a nuclear track neutron detector having at the moment a spatial resolution of about 3 to 4 µm [18]. The nuclear track detectors is made out of CR39 plastic coated with 5mg/cm^2 $^{235}UF_4$. Nuclear fission converts a neutron into a detectable track on CR39. The tracks can be visualized with a standard optical microscope after chemical treatment. The typical diameter of such a track is around 1.5 µm with a length of about 10 µm. Competing reactions from γ rays or alpha particles have a smaller track signature and as a consequence, background from these reactions is practically zero. The automatic readout of the CR39 detector was done in the CHORUS group at CERN. [18]. The microscope MICOS2 is normally used to scan radiated emulsion plates in a search for neutrino oscillation. The rectangular stage of the microscope can be moved by step motors with a reproducibility of 1 µm. The focal length of the microscope is adapted to a CCD camera. The resolution in terms of one pixel is approximately 0.34 µm. An image analysis program detects the tracks on CR39. Having followed the tracks in depth of CR39, the impact point of the fission product on CR39 was found and the spatial resolution of the detector was significantly improved.

The population of the ground state and lowest state follows the quantum mechanical prediction. Higher, unwanted states are removed by the rough neutron absorber made up of an alloy of Ti, Zn and Gd.

4.2 Observation of Quantum States

Signatures of quantum states in the gravitational field of the earth are observed in the following way: A ^{3}He counter measures the total neutron transmission T, when neutrons are traversing the mirror absorber-system as described in Sect. 2. The transmission is measured as a function of the absorber height h and thus as a function of neutron energy since the height acts as a selector for the vertical energy component $E_\perp$, see Fig. 3). The solid data points, plotted in Fig. 5, show the measured number of transmitted neutrons for an absorber height h from zero up to 160 µm. From the classical point of view, the transmission T of neutrons is proportional to the phase space volume allowed by the absorber. It is governed

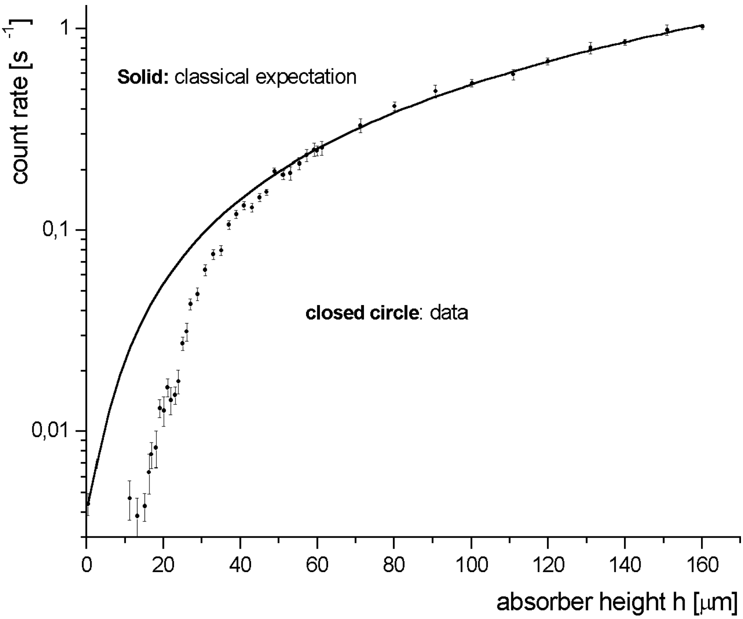


Fig. 5. Data vs. classical expectation

by a power law $T \sim h^n$ and $n = 1.5$. The solid line in Fig. 5 shows this classical expectation.

Above an absorber height of about 60 µm, the measured transmission is in agreement with the classical expectation but below 50 µm, a deviation is clearly visible. From quantum mechanics, we easily understand this behavior: Ideally, we expect a stepwise dependence of T as a function of h. If h is smaller than the spatial width of the lowest quantum state, then T will be zero. When h is equal to the spatial width of the lowest quantum state then T will increase sharply. A further increase in h should not increase T as long as h is smaller than the spatial width of the second quantum state. Then again, T should increase stepwise. At sufficiently high slit width one approaches the classical dependence and the stepwise increase is washed out. Figure 6 shows details of the quantum regime below an absorber height of $h = 50$ µm. The data follow this expectation as described: No neutrons reach the detector below an absorber height of 15 µm as explained before. Then above an absorber height of 15 µm, we expect the transmission of ground-state neutrons resulting in an increase in count rate. The expectation in this case is shown in a solid line and agrees nicely with the data. The χ^2 is 56 for 35 degrees of freedom for one fit parameter, the neutron flux [20]. The expectation for neutrons behaving as classical particles is shown in a dotted line. The classical expectation for neutron transmission is in clear disagreement with the data (open circle). Especially, no neutrons are transmitted for an absorber height between zero and fifteen micrometer.

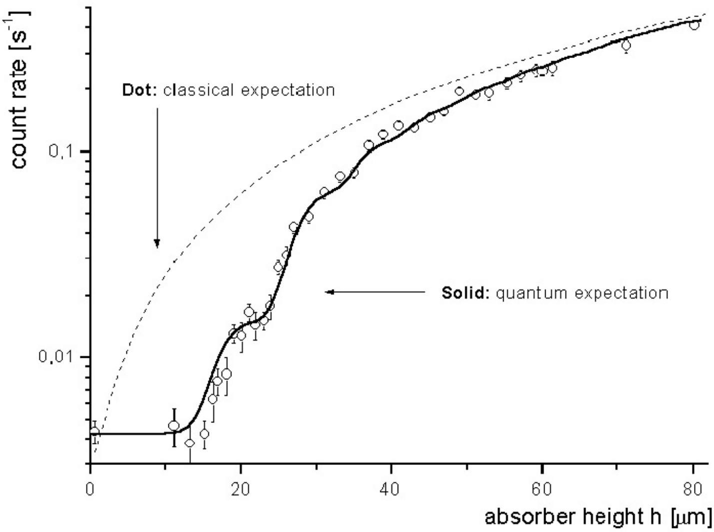


Fig. 6. Data and quantum expectation

5 Summary

In this experiment, gravitational bound quantum states have been seen for the first time. The experiment shows that, under certain conditions, neutrons do not follow the classical Galileian expectation when reflected from neutron mirrors. The measurement does well agree with a simple quantum mechanical description of quantum states in the earth's gravitational field together with a mirror-absorber system. We conclude that the measurement is in agreement with a population of quantum mechanical modes. Further, the spectrometer operates on an energy scale of pico-eVs and suitably prepared mirrors can usefully be employed in measurements of fundamental constants or in a search for non-Newtonian gravity. The present data constrain Yukawa-like effects in the range between 1 µm and 10 µm. This work has been funded in part by the German Federal Ministry (BMBF) under contract number 06 HD 854 I and by INTAS under contract number 99-705.

References

1. V. Nesvizhevsky et al.: *Nature* **415**, 297 (2002)
2. N. Arkani-Hamed, S. Dimpoulos, G. Dvali: *Phys. Lett. B* **429**, 263 (1998)
3. N. Arkani-Hamed, S. Dimpoulos, G. Dvali: *Phys. Rev. D* **59**, 086004 (1999)
4. I. Antoniadis: *Physics with large extra dimensions and non-Newtonian gravity at sub-mm distances*, this volume
5. V.I. Luschikov and A.I. Frank: *JETP Lett.* **28**, 559 (1978)
6. H. Wallis et. al.: *Appl. Phys. B* **54**, 407 (1992)

7. C.G.Aminoff, A.M. Steane, P. Bouyer, P. Desbiolles, J. Dalibard, and Cohen-Tannoudji: *Phys. Rev. Lett.* **71**, 3083 (1993)
8. M.A. Kasevich, D.S. Weiss, and S. Chu: *Opt. Commun.* **15**, 607 (1990)
9. T.M. Roach et al.: *Phy. Rev. Lett.* **75**, 629 (1995)
10. E. Fischbach and C.L. Talmadge: *The Search for Non-Newtonian Gravity*, Springer-Verlag New York 1999
11. S. Lamoreaux: *Phy. Rev. Lett.* **78**, 5 (1997)
12. J.K. Hoskins et al.: *Phys. Rev. D* **32**, 3084 (1985)
13. V.P Mitrofanov and O.I. Ponomareva: *Sov. Phys. JETP* 67 1963 (1988)
14. C.D. Hoyle et al.: *Phys. Rev. Lett.* **86**, 1418 (2001)
15. J. Gundlach: *Equivalence principle tests and tests of the* $1/r$*-law at short distances*, 271^{th} WE-Heraeus-Seminar, 25.2.–1.3. 2002 Bad Honnef
16. A. Westphal: raport de stage, Institut Laue-Langevin, 1999
17. V. Nesvizhevsky et al.: *Nucl. Instr. and Meth. A* **440**, 754 (2000)
18. F. Rueß: *Quantum States in the Gravitational Field*, Diploma thesis, University of Heidelberg, (2000), unpublished
19. I.I. Gurevich and P.E. Nemirovskii: *J. Exptl. Theoret. Phys.* (U.S.S.R.) **41**, 1175 (1961)
20. A. Westphal: Quantum Mechanics and Gravitation, Diploma thesis, University of Heidelberg, (2001), arXiv: gr-qc/0208062

The Einstein Equivalence Principle and the Search for New Physics

Claus Lämmerzahl

ZARM (Center of Applied Space Technology and Microgravity), University Bremen, Am Fallturm, 28359 Bremen, Germany

Abstract. One strategy for searching for effects due to quantum gravity (QG) is to employ the Einstein Equivalence Principle (EEP) as an operational tool. The EEP, consisting of the Universality of Free Fall, the Universality of the Gravitational Redshift, and Local Lorentz Invariance, implies on the one hand that gravity has to be described by a space-time metric (pseudo-Riemannian Geometry) and on the other hand that the equations of motion for particles or quantum fields possess a specific structure compatible with the metrical structure of gravity. Therefore, any deviation from that structure of the dynamical equations of matter is related to a deviation from the metrical structure of gravity. As one consequence, any "new physics" will be accompanied by a breakdown of the validity of the EEP. We now take the EEP as guiding principle for the search for possible QG effects. Various proposed QG effects are then classified according to the various violations of the EEP, which is then used for a comparison with the present experimental status.

1 Introduction

There are many reasons that suggest that gravity has to be quantized [1]. One reason is that if all matter fields are quantized then all fields that are produced by these matter fields have to be quantized, too. Otherwise, at least within the existing theoretical schemes, inconsistencies may occur. Further reasons are that within purely classical General Relativity (GR) under very general and physically plausible circumstances singularities will occur. One way to avoid singularities may be a quantum description of gravity. For completeness one should mention that despite of these reasons there is also the option that gravity and the quantum domain remain completely disconnected so that gravity does not need to be quantized. However, the general accepted opinion is that gravity and quantum theory should go through some synthesis.

There are two main directions along which the quantization of gravity is looked for: The canonical quantization of gravity and string theory. In a sense, both approaches are complementary. The canonical quantization scheme starts from the geometric view of gravity and tries to quantize the gravitational field in form of the space-time metric or other related variables. During this process, the specific properties of matter are completely ignored, as is the case in Einstein's field equations, where matter is summarized in an energy-momentum tensor without specifying the nature of the existing matter. String theory, on the other side, starts from a specific unified concept of particles and interactions in flat

C. Lämmerzahl, The Einstein Equivalence Principle and the Search for New Physics, Lect. Notes Phys. **631**, 367–400 (2003)
http://www.springerlink.com/

space-time, and adds the gravitational interaction which then may depend on the particles it acts on.

It has also been discussed whether a quantization of the gravitational field is sufficient. Perhaps also the notion of an event, that is, a point in the differentiable manifold, should be quantized [2,3]. One approach in this direction is the ansatz of a non-commutative geometry.

However, as far as the experimental search for possible quantum gravity effects is concerned, the effect one is searching for can be described within the ordinary conceptual framework of fields/particles on a differential manifold. Any "new physics" then will appear as a deviation from the usual physical laws. This is obvious because in a low energy limit the usual physical concepts should come out, as, for example, for small velocities the Galilei group results from the Lorentz group or Newtonian gravity is the weak field and low velocity limit of Einstein's theory of gravity. Consequently, any new effect in these examples comes in first by small deviations from the Galilei group or from Newton's gravitational theory.

It has been shown that the low energy limit of QG theories always lead to small deviations from standard physics, mainly due to the appearance of extra scalar fields that are dynamical. These fields couple to the "bare" coupling constants like Newton's gravitational constant G or the fine structure constant α. This makes these constants effectively dynamical, that is, time and position dependent. Furthermore, Local Lorentz Invariance is found to be violated.

After having recognized that QG predicts some deviations from standard physics the next question is how to characterize 'standard physics'. In our context, standard physics is characterized by the Einstein Equivalence Principle (EEP). The EEP first implies that gravity has to be described by a pseudo-Riemannian geometry (gravity is a metric theory) and second gives a formal frame for the description all matter fields and interactions. By specifying a matter field, the structure of the corresponding field equation follows from the EEP. Furthermore, the EEP is stated in terms that are directly related to experimental experience. Therefore, the EEP is a tool to derive essential features of standard physics and serves as an *operational* guiding principle in the experimental search for new physics.

We also want to emphasize that the EEP, though being very fundamental for the general construction of theories, also bears importance for daily life, Fig. 1. The validity of the EEP is directly related to metrology, that is, for example, to the uniqueness of time-keeping or the uniqueness of the definition of other physical units. Also for the Global Positioning System GPS, neglecting GR or Special Relativity (SR) might give daily errors of the order of 10 km. And the high precision observation of the motion of the surface of the Earth with an accuracy of cm, which is at the limit of the present confirmation of SR and GR, can help in the modelling of the Earth and in predicting e.g. Earthquakes.

In this contribution we first describe the EEP, then show most of its implications, present models that lead to violations of the EEP, and give a list of tests of the EEP. At the end we expand a bit the importance of the EEP for metrology, that is, for the task to prepare, reproduce and compare physical units.

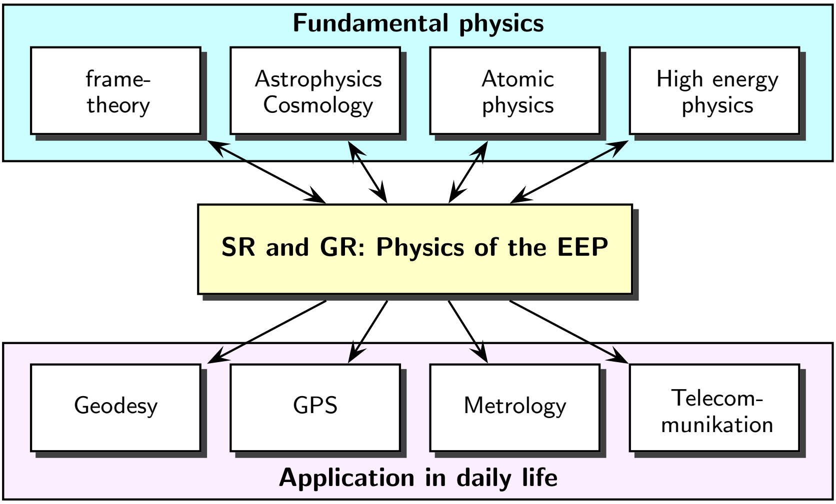


Fig. 1. The importance of GR and SR in fundamental physics and in daily life or, equivalently, the possible influence of "new physics"

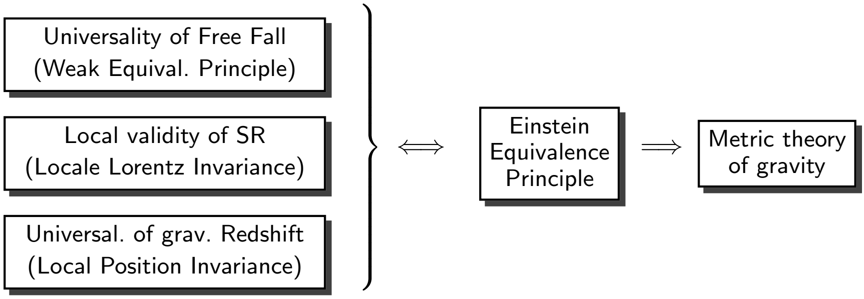
2 The Einstein Equivalence Principle

The EEP is a collection of principles that results in the present day formulation of relativistic physics including SR and GR as well as the Maxwell and the Dirac equation. That means, the validity of the EEP implies the validity of SR, the metrical structure of GR and the form of the equations for the electromagnetic field and for spin-$\frac{1}{2}$ quantum particles. However, the EEP is not enough to derive Einstein's field equations. For that, more input is needed, like the Strong Equivalence Principle. A scheme of how to arrive at Einstein's field equations within a metrical framework is provided by the PPN-formalism [4].

To be more precise, the EEP consists of [4]

1. The principle of the Universality of Free Fall (UFF), which means that all pointlike, structureless particles fall in a gravitational field along the same path,
2. the principle of Local Lorentz Invariance (LLI) which means that in small regions (the region must be small enough so that tidal effects can be neglected with respect to the effects under consideration) SR is valid[1], and
3. the principle of the Universality of the Gravitational Redshift (UGR) which means that all experiments, prepared with the same initial and boundary conditions, give the same results irrespective of where and when they are carried out.

[1] There is a huge set of publications on a precise mathematical and operational meaning of this point, see e.g. [5] for an early reference.



The important point of the EEP is, that it is expressed directly in terms of experimentally testable statements. Thus, it is an operational principle. That means, only if a certain set of experiments yields specific results, then gravity has to be a metric theory, the equations for the electromagnetic field must be of Maxwellian form, the Dirac equation must have their standard form, etc.

The tests of the EEP, and their corresponding meaning, are:

1. Tests of the UFF:
   By testing the UFF one explores whether all constituents of a macroscopic body, that is protons, electrons, neutrons (or the underlying quarks), that is, all forms of rest masses, behave in a gravitational field in the same way. It is certainly an astonishing physical fact that all particles "know" how other particles behave in the gravitational field. In principle, these tests must be carried through for all materials. Since, due to $E = mc^2$, the electromagnetic, weak, and strong interactions also contribute to the rest mass, UFF also controls the behaviour of these interactions in gravitational fields. In a certain theoretical frame these contributions are in general smaller than "pure" violations of the UFF due to e.g. an anomalous gravitational masses.
2. Tests of the UGR:
   These tests explore whether all kinds of clocks based on non-gravitational physics (gravitationally driven periodic systems like the motion of planets around the sun are excluded) possess constant mutual frequency ratios irrespective of their position and time. Since the gravitational field may be different for different space-time points, this means again that all interaction between particles behave in the same way under changes of the gravitational fields. Because a violation of the UFF by the participating particles also would destroy the validity of UGR, UFF and UGR are deeply linked. However, in a first approximation, UFF is connected with the rest mass and UGR with the interactions.
   A (hypothetical) violation of the UGR would mean that the physical laws depend on the time and the position of the laboratory. As an example, assume that the strength of the electric force between two charges depends on time. Since a force is always measured by comparison with another force (or inter-

action), which usually is defined by electromagnetism and quantum mechanics, this means that the electromagnetic coupling constant, the fine structure constant, depends on time. This, in turn, then leads to time-dependent frequency ratios between various clocks, like resonators and atomic clocks. Consequently, the temporal or spatial variation of physical constants is deeply connected with a violation of UGR.

Furthermore, if a violation of the UFF is due to a position-dependent scalar function, then this can be related to a violation of the UGR [6]. That means, UGR-tests are also tests of UFF, and vice versa. General arguments then give that the precision of the determination of the gravitational redshift must be 10^{-10} in order to compete with current UFF tests. See also [7,8] for general considerations of connections of UFF and UGR within string theory inspired dilaton models. In [9] a connection between a varying fine structure constant and violations of UFF is outlined.

3. Tests of LLI

   In order to experimentally verify SR one has to carry through the following set of experiments:

   (a) Test of the universality of the limiting speed of all particles. This includes that all particles possess as maximum speed the speed of light. Only if all phenomena possess the same limiting velocity, causality can be geometrized. That this is the case is again highly non-trivial. As for the UFF and UGR, all particles "know" about a specific property of all other particles.
   (b) Test of the independence of the speed of light from the velocity of the source. As a consequence, this is then also valid for the limiting velocities of all particles.
   (c) Test of the isotropy of the speed of light.
   (d) Tests of the independence of speed of light from the velocity of the laboratory.
   (e) Test of the time dilation given by the Lorentz factor $\gamma = 1/\sqrt{1-v^2}$.

All the above described tests are either direct comparisons between two particles or between two clocks: Tests of the UFF, of the universality of the speed of light, and of the independence of the speed of light from the velocity of the source just compare the velocity of different particles. Since these comparisons are carried through at same space-time events, there is no need for synchronization or for transporting some physical units. These comparisons are null tests and are completely *independent* from any theoretical model or framework.

The second set of tests, namely test of the UGR, of the isotropy of the speed of light, of the independence of the speed of light from the velocity of the laboratory, and of the relativistic time dilation are comparisons between different clocks. The first three of these tests compare two clocks of different physical nature in the same state of motion at the same position, see Fig. 2. The last compares two identical clocks possessing different velocities.

Though the comparison of clocks means that one just measures the ratio of two frequencies that is independent of any time unit, the description and

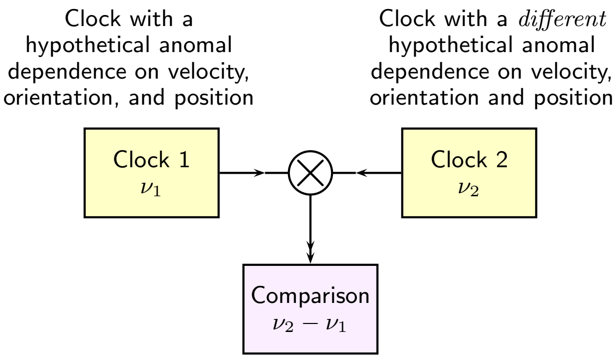


Fig. 2. General scheme for testing parts of SR and GR with clocks, namely the UGR, isotropy and velocity independence of the speed of light

experimental maintenance of clocks needs theoretical and experimental effort. Clocks based on resonators, for example, need, if high precision is required, very careful thermal and mechanical stabilization. In the case of atomic clocks, external stray fields have to be under very precise control. That means, that one has to use already some laws of physics in order to prepare these clocks. Only with a careful order-of-magnitude analysis of the physics needed for establishing the clocks compared to the laws one is testing, one can avoid logical inconsistencies. Furthermore, if clocks are used for searching for anomalous effects violating the EEP, the result can be interpreted consistently only if also the clock is described as a whole within this EEP-violating model.

Furthermore, from the available high-precision clocks one can infer the kind of information one may get from the corresponding clock-comparison experiments. Such clocks are

1. Atomic clocks based on electronic hyperfine transitions. They are characterized by energy levels of the form $E = \alpha^2 f(\alpha)$ where α is the fine structure constant and where $f(\alpha)$ is a Casimir correction factor and depends on the corresponding transition. An external field needed in order to split energy levels.
2. Atomic clocks based on nuclear transitions. These "clocks" are not used in practice as primary clocks. However, these transitions are used e.g. in Hughes–Drever experiments in order to search for LLI violations. The transitions are characterized by a nuclear fine structure constant $\alpha_{\rm n}$. Again one has to apply an external magnetic field.
3. Resonators. Here the frequency is defined by microwave or optical frequencies in a resonator. Since the length of the resonator is given by Bohr's radius, it scales linearly with the fine structure constant. Therefore, the frequencies possess a different α-dependence than atomic clocks. A direct comparison gives information about a hypothetical time-dependence of α which violates UGR [75].

4. Molecular clocks. The rotation or vibration of molecules also define a frequency which depends on the fine structure constant but also on the ratio of the electron-to-proton-mass: $E \sim f(\alpha, m_e/m_p)$.

Consequently, by comparison of these various clocks one may get information about the constancy of α, α_{n} and m_e/m_p. Furthermore, since in all these clocks characteristic directions are involved (atomic clocks need some external electric or magnetic field, the geometry of resonators in general possess a symmetry axis, and molecules possess some axis of rotation or direction of vibration) a comparison of clocks with different intrinsic directions represent also tests of the rotation invariance of physics (tests of the isotropy of space, of which Michelson–Morley tests are the realization in the electromagnetic domain).

3 Implications of the Einstein Equivalence Principle

All interactions are discovered and characterized by the influence on matter. This is also true for the gravitational interaction. The EEP consists of statements about the properties of matter in gravitational fields from which we can conclude the properties of that field. That means, the EEP first very strictly restricts the equations for the electromagnetic field and for point particles or quantum fields. It is only through the specific properties of the dynamics of matter (point particles or fields) that gravity can be restricted to be describable solely by means of a space-time metric. Therefore, we first have to analyze the dynamics of particles and fields prescribed by the EEP and then we can define the gravitational field as that field which couples universally to matter and deduce the properties of this gravitational field.

Roughly speaking, the UFF implies the geometrization of the gravitational interaction since no particle properties influence the dynamics, LLI implies the existence of a metric tensor at each space-time point, and, finally, UGR implies that there are no scalar or other fields leading to different metrics at different space-time points.

3.1 Matter

As already stated, the EEP not only restricts the structure of the gravitational field but in a first step the structure of dynamical equations like the equation of motion for point particles, for quantum fields or of the electromagnetic field. For example no coupling to the curvature is allowed because curvature terms will not vanish when restricting the experiment to small regions where the dynamical equations are assumed to acquire their SR form. It should be emphasized that this holds only for the equations governing *observed* quantities, like the electromagnetic field. It is well known from Maxwell equations minimally coupled to gravity that the equations governing the vector potential couples to the curvature. The same also appears in field theory: Here the requirement that the fundamental solution of a scalar field equation has the same form as in SR leads to a conformal coupling, that is, to a curvature term in the field equation [10].

Point Particles. If we accept that in any situation the position and velocity is enough to determine the path of a structureless particle, then the equation of motion is, in general, given by $\ddot{x}^\mu + \widetilde{H}^\mu(p, x, \dot{x}) = 0$, where p denotes all effective parameters (charge-to-mass ratio q/m, deviation of the ratio of the gravitational to inertial mass from unity $m_\mathrm{g}/m_\mathrm{i} - 1$, etc.) characterizing the particle under consideration[2]. By taking $p \to 0$ (either in a continuous or discrete way and leaving the masses finite) we may define that part of $\widetilde{H}$ which is independent of p, $H^\mu(x, \dot{x}) = \lim_{p\to 0} \widetilde{H}^\mu(p, x, \dot{x})$. Then the equation of motion may be written as $\ddot{x}^\mu + H^\mu(x, \dot{x}) + \widehat{H}^\mu(p, x, \dot{x}) = 0$ with $\widehat{H}^\mu = \widetilde{H}^\mu - H^\mu$. That part which is independent of any parameters p we identify as *gravitational interaction.* $\widehat{H}^\mu$ is identified with nongravitational interactions. These terms are not present either if the corresponding charges are zero (neutral particles) or if there is no non-gravitational field. In this case we have, per definition, UFF. We used the UFF as a means to identify the gravitational interaction.

In a second step, LLI introduces at each point a frame with a Minkowskian metric. This is physically represented by the motion of light rays (up to a conformal factor).

Now we further consider structureless particles which are either neutral or which move in an interaction-free region. The equation of motion then must have the form $\ddot{x}^\mu = H^\mu(x, \dot{x})$ with no parameter noticing any property of the particle. According to LLI, there is a frame so that the equation of motion acquires the SR form $\ddot{x}^\mu \overset{*}{=} 0$. In this frame we also have the Minkowski metric. There cannot be any coupling of the particles to curvature like $R^\mu{}_\nu \dot{x}^\nu$ because such terms are present in any frame. The transformation of $\ddot{x}^\mu = 0$ to an arbitrary frame then yields an autoparallel equation $v^\nu \partial_\nu v^\mu + \Gamma^\mu_{\nu\rho} v^\nu v^\rho = \alpha v^\mu$ where α is an undetermined function. Furthermore, the Minkowski metric will transform to a metrical tensor $g_{\mu\nu}$. The compatibility of this autoparalell with SR then leads to a Weylian connection [11,12]. Furthermore, the condition of UGR in the form of a uniqueness of a transport of light clocks then reduces the Weylian connection to a Riemannian one.

Consequently, for point particles the EEP reduces all possible gravitational interactions to the one described by a Riemannian geometry.

Matter Fields. Also the standard Dirac equation can be derived with the help of the EEP. Assuming the conservation of probability, the requirements of LLI leads to a system of first order partial differential equations, which have the form of a slightly generalized Dirac equation. Adding the principle of UGR one then arrives at the usual Dirac equation in pseudo-Riemannian space-time [13].

The Maxwell Field. It has been shown by Ni [14] that only a modest generalization of Maxwell's equations is compatible with the UFF. The modification

[2] We exclude non-scalar properties of particles because this will considerably complicate the procedure because then one has to take the dynamics of these properties into account which increases the equations of motion under consideration.

consists of an addition of $\chi\epsilon^{abcd}F_{ab}F_{cd}$ to the usual Lagrangian for the Maxwell field where χ is a pseudoscalar field and ϵ^{abcd} the totally antisymmetric Levi–Civita tensor[3]. However, this term violates LLI because it induces a precession of the polarization of plane waves. This indeed represents a violation of LLI because corresponding propagation phenomena induced by the propagation of light do not show such a precession. Requiring LLI then forbids such a precession and thus this extra term. Therefore, EEP implies the ordinary Maxwell equations.

3.2 The Gravitational Field

The gravitational field is now defined through the equations of motion of the various matter fields. The only gravitational interaction, which remain in the dynamical equations of point particles, the Maxwell and the Dirac field, is given by a space-time metric. This metric then is called the gravitational field. Gravity then is described mathematically by a pseudo-Riemannian geometry [4].

4 Models Which Violate the Einstein Equivalence Principle

4.1 Quantum Gravity Induced Violations of the EEP

String Theory Induced Violations of the UFF. The prediction [17,18] is that the UFF, in terms of the Eötvös parameter (20) below, might be violated at the order 10^{-15} or even at the order 10^{-13} [19]. This is very well in the range of the space mission MICROSCOPE [20] scheduled for 2006.

Loop Gravity Induced Violations of LLI.

Modifications of the Maxwell Equations. In loop gravity, averaging over some quasiclassical quantum state, a so-called "weave"-state, which includes the state of the geometry as well as of the electromagnetic field, gives the effective Maxwell equations [21] (see also [22])

$$
\begin{aligned}
0 &= \boldsymbol{\nabla} \times \boldsymbol{B} - \partial_t \boldsymbol{E} + \vartheta_1 \boldsymbol{\nabla} \times \boldsymbol{B} \\
&\qquad + \vartheta_2 \Delta(\boldsymbol{\nabla} \times \boldsymbol{B}) + \vartheta_3 \Delta \boldsymbol{B} + \vartheta_4 \boldsymbol{\nabla} \times (B^2 \boldsymbol{B}) + \ldots \qquad (1) \\
0 &= \boldsymbol{\nabla} \times \boldsymbol{E} + \partial_t \boldsymbol{B} + \vartheta_1 \boldsymbol{\nabla} \times \boldsymbol{E} + \vartheta_2 \Delta(\boldsymbol{\nabla} \times \boldsymbol{E}) + \vartheta_3 \Delta \boldsymbol{E} + \ldots \,, \qquad (2)
\end{aligned}
$$

[3] One should distinguish between Lorentz invariance and Lorentz covariance: Lorentz covariance means the formal covariance of all expressions under Lorentz transformation, Lorentz invariance means that the result of all experiments are the same in all frames provided all initial and boundary conditions in each frame are identical. Therefore, a Lorentz covariant theory may break Lorentz invariance. As an example, applying an equivalence principle of the form that e.g. the Dirac equation should acquire its special relativistic form in a particular frame, allows a coupling to space-time torsion [15,16] what clearly introduces distinguished space-time directions.

where the ϑ_i are coefficients depending on ratios of the Planck length and a length characterizing the quasiclassical gravitational quantum state. If one introduces the plane wave ansatz $\boldsymbol{E} = \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$ and $\boldsymbol{B} = \boldsymbol{B}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$ and neglects the nonlinearities, then one gets the dispersion relations

$$\omega = |\boldsymbol{k}| \left(1 + \tilde{\theta}_1 + \tilde{\theta}_2 |\boldsymbol{k}|^2 \pm \tilde{\theta}_3 |\boldsymbol{k}|\right) \tag{3}$$

with other $\tilde{\theta}$s. The $\pm$ corresponds to different polarization states. Moreover, dispersion occurs so that from this dispersion relation one can e.g. derive the group velocity of photons and analyze the time of arrival of signals from distant stars according to frequency or polarization.

There are two points which need to be discussed: First, with (2) also a homogeneous Maxwell equation is modified and, second, the appearance of higher order derivatives means that there are photons that propagate with infinite velocity.

1. The deviation from the homogeneous Maxwell equations has the important consequence that the unique description of charged particle interference is no longer true. If quantum mechanical equations are minimally coupled to the electromagnetic potential, that is by the replacement $\partial \to \partial - \frac{ie}{\hbar c} A$, then the phase shift in charged particle interferometry is $\delta\phi = \frac{ie}{\hbar c} \oint_C A$ for a closed path C. If one applies Stokes' law in the case of a trivial space-time topology in that region, then this is equivalent to $\delta\phi = \frac{ie}{\hbar c} \int F$ with $F = dA$ where integration is over some 2-dimensional surface bounded by the closed path C. The fact that the result should not depend on the chosen surface is secured by the homogeneous Maxwell equations $dF = 0$.
   Since the form of (2) is incompatiple with dF=0, one must consider nonminimal couplings in order to provide unique predictions for charged particle interference.
2. The appearance of higher order spatial derivatives implies that in the corresponding frame of reference there are photons propagating with infinite velocity. As a consequence, Lorentz-invariance is violated.
3. It is clear that the appearance of higher order spatial derivatives also implies the appearance of higher order time derivatives in other frames of reference.

Modifications of the Dirac Equation. In the same manner as for the Maxwell equation, one can derive the modified Dirac equation [23,24]. The effective Dirac equation has the form

$$i\widetilde{\gamma}^a \partial_a \psi - \widetilde{m}\psi - \widetilde{\gamma}^{ab} \partial_a \partial_b \psi = 0\,, \tag{4}$$

where $\widetilde{\gamma}^a = \gamma^a + \kappa_1 G_1 (L_{\rm PL}/L) + \kappa_2 G_2 (L_{\rm PL}/L)^2 + \ldots$ are the usual Dirac matrices γ^a with QG corrections (κ_i are coefficients of order 1, G_i are some matrices, and $L_{\rm Pl}$ and L are the Planck length and a length characterizing the semiclassical gravitational quantum state, respectively), and $\widetilde{m} = m\big(1 + \mu_1 m (L_{\rm PL}/L) + \mu_2 (L_{\rm PL}/L)^2 + \ldots\big)$, and $\widetilde{\gamma}^{ab} = \gamma^{ab}\left(\lambda_1 (L_{\rm PL}/L) + \lambda_2 (L_{\rm PL}/L)^2 + \ldots\right)$ where again μ_i and λ_i are parameters of order unity, and γ^{ab} is some set of matrices.

This kind of equation can be used in order to discuss the time-of-arrival of neutrons [23,25] or can be confronted with Hughes–Derever experiments [26,27]. The result of the latter paper is that modifications linear in the Planck length are questionably.

String Theory Induced Violations of LLI.

Modifications of Maxwell Equations. In string theory the gravitational field is given by D-branes which interact with propagating photons via an effective recoil velocity $\bar{\boldsymbol{u}}$ [28]. This recoil effect appears as a modified space-time metric, which influences the Maxwell equations. These equations are then given by

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}+\bar{\boldsymbol{u}}\cdot\partial_t\boldsymbol{E}=0 \qquad \boldsymbol{\nabla}\cdot\boldsymbol{B}=0 \tag{5}$$

$$\boldsymbol{\nabla}\times\boldsymbol{B}-(1-\bar{u}^2)\partial_t\boldsymbol{E}+\bar{\boldsymbol{u}}\times\partial_t\boldsymbol{B}+(\bar{\boldsymbol{u}}\cdot\boldsymbol{\nabla})\boldsymbol{E}=\;0 \qquad \boldsymbol{\nabla}\times\boldsymbol{E}=-\partial_t\boldsymbol{B}\,. \tag{6}$$

In this approach the homogeneous equations remain unmodified. The resulting wave equations are

$$0=\Box\boldsymbol{E}-2(\bar{\boldsymbol{u}}\cdot\boldsymbol{\nabla})\partial_t\boldsymbol{E} \tag{7}$$

$$0=\Box\boldsymbol{B}-2(\bar{\boldsymbol{u}}\cdot\boldsymbol{\nabla})\partial_t\boldsymbol{B} \tag{8}$$

which leads to the dispersion relation

$$\omega=\pm k+(\bar{\boldsymbol{u}}\cdot\boldsymbol{k})+\mathcal{O}(\bar{u}^2)\,. \tag{9}$$

The corresponding group velocity for light is

$$\boldsymbol{c}=\pm\frac{\boldsymbol{k}}{k}+\bar{\boldsymbol{u}}\,. \tag{10}$$

From string theoretical considerations, the recoil velocity $\bar{\boldsymbol{u}}$ can be shown to depend linearly on the energy ω of the photon, $\bar{\boldsymbol{u}}\sim\omega$. Therefore, in this case we obtain an energy dependent group velocity of the photons.

Modifications of the Dirac Equation. String theory induced modifications of the effective Dirac equation [29] have the form

$$i\gamma^a\partial_a\psi-\bar{u}^a\gamma^0 i\partial_a\psi-m\psi=0\,. \tag{11}$$

Also this equation can be confronted with spectroscopic results [30] with the result that first order corrections coming out from (11) are unlikely to be present.

4.2 LLI Violations from Non-commutative Geometry

It has been shown [31] that non-commuative geometry in general leads to violations of LLI described in general by an extension of the standard model (see below). In a non-commutative framework the commutator of coordinates x^μ is

$[x^\mu, x^\nu] = i\theta^{\mu\nu}$ where $\theta^{\mu\nu}$ is real and antisymmetric. The main argument employed in [31] is that due to the Seiberg–Witten map stating that there is a correspondence between a non-commutative gauge theory and a conventional gauge theory, non-commutative models must lie within an extension of the standard model. They applied the Seiberg–Witten map to the model of non-commutative QED and, after restricting to quadratic terms, arrived at an effective Lagrangian, which is within the model given by sum of (16) and (19) together with a coupling term that we will describe below. The violations of LLI are directly related to the coefficients $\theta^{\mu\nu}$.

The General Structure of Modified Dispersion Relations. In both cases, that is in (3) and (9), the general structure of dispersion relation for the propagation in one direction reads [32]

$$\boldsymbol{k}^2 c^2 = E^2 \left(1 + \xi \left(\frac{E}{E_{\mathsf{QG}}}\right) + \mathcal{O}\left(\frac{E}{E_{\mathsf{QG}}}\right)^2\right) . \tag{12}$$

where E is the energy of the photon. The parameter ξ depends on the underlying theory and can be derived to be $\sim 3/2$ for string theory [33] and ~ 4 for loop gravity [22]. The quantum gravity energy scale E_{QG} is, of course, of the order of the Planck energy E_P so that for ordinary light which possesses an energy of the order 1 eV the correction term E/E_{QG} is of the order 10^{-28}. From the above dispersion relation we derive the velocity of light

$$c_{\mathsf{QG}} = c \left(1 - \xi \frac{E}{E_{\mathsf{QG}}}\right) . \tag{13}$$

Therefore, the difference of the velocity of light for high energy photons and low energy photons, $\varDelta c = c(E) - c(E \to 0)$ is given by

$$\frac{\varDelta c}{c} = \xi \frac{E}{E_{\mathsf{QG}}} . \tag{14}$$

Exactly this quantity has been suggested to be observed for astrophysical events.

String Theory Induced Violation of the UGR. Violations of the UGR induced by string theory have been considered within a dilaton model by Damour [7,8].

4.3 General Models Violating EEP

Bi-metric Models. Bi-metric models describe the possibility that the limiting velocities of different kinds of particles may differ. Thus they are test theories for the universality of the velocity of light. If the speed of light is not universal then is also has as consequence that the isotropy of the speed of light and its independence from the velocity of the laboratory also will be violated.

The most elaborate model of this kind is the TH$\epsilon\mu$-formalism [34,35,4]. It is based on the Lagrangian

$$S = m_0 \int \sqrt{T - H\dot{\boldsymbol{x}}^2} dt + \frac{1}{8\pi} \int \left(\epsilon \boldsymbol{E}^2 - \frac{1}{\mu} \boldsymbol{B}^2 \right) d^4x + q \int \boldsymbol{A} \cdot \dot{\boldsymbol{x}} dt + q \int A_0 dt \,, \tag{15}$$

where $\dot{\boldsymbol{x}}$ is the coordinate velocity of a point particle with mass m and charge q, $\boldsymbol{E}$ and $\boldsymbol{B}$ the electric and magnetic field, A_0 and $\boldsymbol{A}$ the scalar and vector potential. T, H, ϵ, and μ are the parameters of this theory which are all unity in standard physics.

From a quick sight at this equations it is clear that the limiting velocity of the point particles is $c_{\mathrm{p}} = \sqrt{T/H}$ while the velocity of photons is $c_{\mathrm{em}} = \sqrt{\epsilon\mu}$. Only if the mechanical parameters T and H are related to the electromagnetic parameters ϵ and μ, both velocities are equal. It is also clear that the Lagrangian (15) is written in a preferred frame characterized by the isotropy of both limiting velocities. In moving frames the difference between these two velocities depends on the direction.

For constant parameters T, H, ϵ, and μ this is a one-parameter test theory for describing tests of SR. This parameter is $\delta = c_{\mathrm{p}}/c_{\mathrm{em}} = \sqrt{T/(H\epsilon\mu)}$. By replacing the point particle part in the Lagrangian by a corresponding Dirac–Lagrangian a lot of experiments including tests of UGR and LLI can be described [36,35,37,38,4].

Ni–Haugan–Kostelecky-Formalism. A huge generalization of the $TH\epsilon\mu$-ansatz consists in the consideration of general constitutive laws between the electromagnetic field strengths $\boldsymbol{E}$ and $\boldsymbol{B}$ and the electromagnetic excitations $\boldsymbol{D}$ and $\boldsymbol{H}$. To the knowledge of the author, Ni [39,14,40] was the first who considered this as a general starting point for the confrontation of the consequences with observations. The starting point is to replace the Lagrange density of the electromagnetic field according to

$$\mathcal{L} = -\frac{1}{16\pi} \eta^{ac} \eta^{bd} F_{ab} F_{cd} \quad \longrightarrow \quad \mathcal{L} = -\frac{1}{16\pi} \left(\eta^{ac} \eta^{bd} + \lambda^{abcd} \right) F_{ab} F_{cd} \,, \tag{16}$$

where λ^{abcd} is an additional tensor possessing the symmetries of the Riemann-tensor. This tensor is assumed to describe the properties of the vacuum. The homogeneous Maxwell equations are still valid. Due to this, the totally anti-symmetric part of λ^{abcd} transforms into total divergence, and the double trace amounts to a rescaling of the charge in the case that there is a coupling to matter, so that there are 19 parameters related to a violation of LLI.

With the general constitutive law $G^{ab} = (\eta^{ac}\eta^{bd} + \lambda^{abcd})F_{cd}$ the inhomogeneous Maxwell equations are

$$\partial_b G^{ab} = 4\pi j^a \,. \tag{17}$$

These equations have been used by Ni [39,14] to derive general conditions for the validity of the UFF for electromagnetic bound systems. Haugan and Kauffmann [41] used this in order to derive constraints on these coefficients from astrophysical observations. The most general analysis of consequences of these LLI-violating terms is due to Kostelecky and Mewes [42,43] who analyzed astrophysical data leading to estimates $\lambda^{abcd} \leq 10^{-30}$ for 10 of the 19 components, and showed how to treat data from laboratory experiments like Michelson–Morley experiments. First analyses of recent laboratory experiments in this test theory have been carried through by Lipa *et al* [44] and Müller and coworkers [45]. While the first paper gives estimates $\lambda^{abcd} \leq 10^{-13}$ resp. 10^{-9} for four linear combinations of the other 9 coefficients, the latter achieved $\lambda^{abcd} \leq 10^{-13}$ resp. 10^{-9} for all individual coefficients but one linear combination.

In more general models one starts with the equation of motion instead of using a variation principle. The most general Maxwell equation linear in the field and the derivative is given by [46,47]

$$(\eta^{ac}\eta^{bd} + \chi^{abcd})\partial_b F_{cd} + \chi^{acd} F_{cd} = 4\pi j^a \,. \tag{18}$$

Here, $\chi^{abcd} = \chi^{ab[cd]}$ possesses 96 and $\chi^{acd} = \chi^{a[cd]}$ 24 degrees of freedom. A first consequence of this general ansatz is that charge conservation is no longer true which has a severe impact in the standard formalism of physics. However, there seems to be no really good test of charge conservation (see below) and, furthermore, recently some papers discussed charge non-conservation models originating in higher dimensional brane theories where the charge may escape from our 4-dimensional world through higher dimensions [48]. Also its relation to the EEP has been discussed [49]. Charge non-conservation is encoded in the coefficients $\chi^{(ab)cd}$ and $\overset{0}{\chi}{}^{acd}$ where in $\overset{0}{\chi}$ the totally antisymmetric part has been removed.

Generalized Dirac Equation. The above-mentioned generalization of the $TH\epsilon\mu$-formalism has also a counterpart on the side of the particles, namely the generalized Dirac equation. Such generalized Dirac equations have the form

$$i\gamma^\mu\partial_\mu\psi - M\psi = 0 \,, \tag{19}$$

where γ^μ are *not* assumed to fulfill a Clifford algebra. A substantial part of LLI-violating effects of this kind of equations is due to the non-Clifford parts of the γ-matrices. Together with the generalized Maxwell equation (17) this is also called the 'extended standard model'. In GR, the matrix M consists of the mass scalar and the spinor connection. In our generalized Dirac equation additional terms will spoil LLI and also UGR.

To the knowledge of the author, generalized Dirac equations and their consequences for particle dynamics have first been discussed by Liebscher [50], also [51]. An early discussion of the generalized Dirac equation with respect to tests of hypothetical violations of LLI is [52], recent discussion are due to Kostelecky and coworkers [53–58].

5 Experimental Tests of the Einstein Equivalence Principle

According to the principles underlying the EEP, the tests of it or, equivalently, the search for new physics can be classified along the following lines:

- Tests of the UFF
- Tests of the UGR
- Tests of LLI
  - Test of the universality of c
  - Test of the independence of the c from the velocity of the source
  - Test of the isotropy of c
  - Test of the independence of c from the velocity of the laboratory
  - Test of time dilation

5.1 Test of the Universality of Free Fall

Usually, tests of the UFF are described within a Newtonian framework: In the system where the gravitating body is at rest, the force acting on a test body, $m_\mathrm{i}\ddot{\boldsymbol{x}}$, where m_i is the inertial mass, is given by the gravitational force $-m_\mathrm{g}\boldsymbol{\nabla}U$, where m_g is the gravitational mass of the test body and U the Newtonian potential. The path of the test body can be determined from the acceleration $-\ddot{\boldsymbol{x}} = (m_\mathrm{g}/m_\mathrm{i})\boldsymbol{\nabla}U$. If the ratio $m_\mathrm{g}/m_\mathrm{i}$ is the same for all test bodies, then the path will also be the same. By renormalizing the Newtonian potential by a constant, we then have $\ddot{\boldsymbol{x}} = -\boldsymbol{\nabla}U$. A hypothetical violation of the UFF is encoded in the Eövös ratio η defined as

$$\eta = 2\frac{\ddot{\boldsymbol{x}}_2 - \ddot{\boldsymbol{x}}_1}{\ddot{\boldsymbol{x}}_2 + \ddot{\boldsymbol{x}}_1} = 2\frac{(m_\mathrm{g}/m_\mathrm{i})_2 - (m_\mathrm{g}/m_\mathrm{i})_1}{(m_\mathrm{g}/m_\mathrm{i})_2 + (m_\mathrm{g}/m_\mathrm{i})_1}\,, \tag{20}$$

where the indices 1 and 2 denote two different test bodies. The UFF implies $\eta = 0$.

The best test gives $\eta \leq 10^{-12}$ [59]. There are two space mission under way, the French MICROSCOPE mission [20] that is scheduled for 2005 but may be delayed for a year due to financial reductions in space programs, and the STEP project [60]. These missions want to test the UFF to a precision of 10^{-15} and 10^{-18}, respectively. In principle, UFF should be tested with *all* pairs of test bodies. Due to new predictions [61] the UFF has recently been probed for small distances [62,63].

5.2 Test of the Universality of the Gravitational Redshift

Absolute Measurement. In GR the gravitational red shift is given by

$$\nu(x_1) = \left(1 - \frac{U(x_1) - U(x_0)}{c^2}\right)\nu(x_0)\,, \tag{21}$$

where $\nu(x_0)$ is the frequency of a clock at position x_0 and $\nu(x_1)$ is the frequency of this clock measured by an identical clock at position x_1. This relation has been tested with a space-borne Hydrogen maser compared with a ground H-maser with an accuracy of $7 \cdot 10^{-5}$ [64,65].

Clock Comparison. In the above formula no reference is made to the used clock. In the case that the gravitational red shift is not universal, the frequencies of the various types of clocks at different positions in the gravitational field will depend on the *type* of the clock:

$$\nu(x_1) = \left(1 - (1 + \alpha_{\text{clock}}) \frac{U(x_1) - U(x_0)}{c^2}\right) \nu(x_0) \,. \tag{22}$$

In the framework of GR, $\alpha_{\text{clock}} = 0$ for all clocks, such as atomic clocks, optical and microwave resonators, H-maser, quartz crystal, etc. If the redshift depends on the type of clock and we move two different clocks together in the gravitational field, then the ratio of the frequencies of these two clocks is

$$\frac{\nu_{\text{clock1}}(x_1)}{\nu_{\text{clock2}}(x_1)} \approx \left(1 - (\alpha_{\text{clock2}} - \alpha_{\text{clock1}}) \frac{U(x_1) - U(x_0)}{c^2}\right) \frac{\nu_{\text{clock1}}(x_0)}{\nu_{\text{clock2}}(x_0)} \,. \tag{23}$$

For a violation of the UGR we get a *position dependent* frequency ratio which is proportional to the difference of the gravitational potential difference $U(x_1) - U(x_0)$. If $\alpha_{\text{clock2}} = \alpha_{\text{clock1}}$, then this frequency ratio is independent of the position of the two clocks, and the constant factor α_{clock} can be absorbed into the Newtonian potential, leading to (21).

The best tests of the UGR have been carried through by comparing an H-maser and a Cs atomic fountain clock over one year. Both are located at the same position on the surface of the Earth and experience, due to the annual elliptical motion of the Earth, the varying gravitational potential of the sun which is of the order $\Delta U/c^2 \sim 7 \cdot 10^{-10}$. The result is $|\alpha_{\text{H-maser}} - \alpha_{\text{fountain}}| \leq 2.1 \cdot 10^{-5}$ [66]. Other tests compare the frequency of a Cesium atomic clock and that defined by a microwave resonator which leads to $|\alpha_{\text{Cs}} - \alpha_{\text{cavity}}| \leq 2 \cdot 10^{-2}$ [67]. A comparison between electronic Iodine states and a cavity yields $|\alpha_{\text{Iod}} - \alpha_{\text{cavity}}| \leq 2 \cdot 10^{-2}$ [68]. The space mission ACES [69] comparing an H-maser and an atomic fountain clock on the ISS in the varying gravitational potential of the sun, aims to improve the presently best test by at least one order what is a consequence of the fact that the free fall condition in space will considerably improve the working conditions of the atomic fountain clock (see below p. 388). Furthermore, since the above results depend on the value of the experienced difference of the gravitational potential, space missions like SPACETIME [70] and OPTIS [71] will give huge improvements. While OPTIS compares an H-maser, atomic ion clocks and clocks based on optical resonators in an high elliptic orbit around the Earth, SPACETIME uses three ion clocks in an identical environment and aims at exploiting the huge potential difference of $\Delta U/c^2 = 3 \cdot 10^{-7}$ when approaching the sun up to 5 solar radii. As a result, UGR may be tested to an accuracy of 10^{-10}.

Since a violation of UGR can be related to charge non-conservation, we report about its status of experimental verification.

Charge Non-conservation. One way to treat charge non-conservation experimentally is to look for the probability that an electron, carrying a charge e, may just disappear or be created during scattering processes in high energy accelerators. The underlying model is that an electron decays into a neutral particle and a photon $e \to n + \gamma$. For these kind of processes the probability was found to be less than $5.3 \cdot 10^{-21}\ \mathrm{y}^{-1}$ [72].

Another aspect of charge conservation is the equality of charges of the electron and the proton. This can be proven with great accuracy by testing the neutrality of atoms or molecules. The corresponding estimates state that the relative difference between the electron and proton charge $|(q_e - q_p)/q_e|$ is less than $2 \cdot 10^{-19}$ [73]. It is clear that even for equal electron and proton charges the absolute charge may vary in time. In principle, this may be proven by observing a spring connecting two, say, equally charged bodies. For a varying, non-conserved charge the spring should expand with time, if the charge decreases. However, since also the physics of the spring heavily depends on the electromagnetic properties of matter, this requires a very thorough analysis. A more simpler version of this can be found in atomic physics where obviously charge non-conservation influences the binding energy of electron in the field of the nucleus what results in a time-dependence of the fine structure constant $\alpha = e^2/\hbar c$.

If we assume that charge is not conserved, then the charge of all particles depends on time so that especially for the charge of the electron and the proton $de/dt \neq 0$ which then implies for the fine structure constant $\dot\alpha/\alpha = 2\alpha(1/e)(de/dt)$. If we assume a specific time-dependence of the form $de/dt = \zeta e$, then $d\alpha/\alpha = 2\zeta$. Thus experiments on the time-dependence on the fine-structure constant give estimates on ζ or, in terms of the general model (18), on components of $\chi^{(ab)cd}$ and $\overset{0}{\chi}{}^{abc}$.

A measurement of a hypothetical time dependence of the fine structure constant can be obtained by comparing different time or length standards which depend in a different way on α (see page 372). For example, the in the recent experiment [74] a Cs and a Rb atomic fountain clock, both based on hyperfine transitions with different α-dependence, were compared over five years. The comparison resulted in $\dot\alpha/\alpha \leq 1.6 \cdot 10^{-15}\ \mathrm{y}^{-1}$, so that $\zeta \leq 8 \cdot 10^{-16}$. In two new proposals [75,76] using resonators, more specific, monolithic resonators for optical modes and whispering gallery modes in a single resonator, respectively, it is claimed that it might be possible to test the time-independence of α below the 10^{-15} level. Together with the tests of the equality of the electron and proton charge, tests of the constancy of the fine structure constant provide the best direct experimental proof of charge conservation.

5.3 Test of Local Lorentz Invariance

While it is clear that for a description of tests of UFF and UGR one has to modify the equations of motion of the test particles and test clocks, the situation is different for SR. Here one may choose between kinematical and dynamical test theories. Kinematical test theories, based on the analysis of Robertson [77] and Mansouri and Sexl [78–80] (see also the review [81]), compare the physics in two differently moving and differently oriented laboratories. Dynamical test theories examine the structure of the laws of physics. Kinematical test theories are more basic in the sense that they describe the behaviour of distinguished physical phenomena (e.g. light) with respect to *given* rods and clocks which, in a constructive approach to a physical theory, at the beginning are given by definition and thus cannot be analyzed using physical theories (simply because they not yet exist at that stage)[4]. In a later stage, after having explored physical laws, one may analyze these rods and clocks with the help of these laws. Then one may ask how these objects behave if these laws are modified. This is the task of the dynamical test theories. In the end, dynamical test theories are superior to kinematical test theories because these theories describe all objects, even the measuring apparatus.

Within the kinematical framework of Robertson [77] and Mansouri and Sexl [78–80] the velocity of light is given by

$$c(v, \vartheta) = c_0 \left(1 + A \frac{v^2}{c_0^2} + B \cos^2 \vartheta \frac{v^2}{c_0^2} + \mathcal{O}\left(\frac{v^4}{c_0^4} \right) \right) , \tag{24}$$

where v is the velocity of the laboratory with respect to a preferred frame which one chooses as the frame in which the cosmic microwave background radiation appears to be isotropic. c_0 is the velocity of light in the preferred frame. A and B are two parameters which vanish in SR. In addition, the time dilation is described as

$$\gamma(v) = \frac{1}{\sqrt{1 - v'^2/c_0^2}} \left(1 + \frac{1}{2} \alpha \frac{(\boldsymbol{v} + \boldsymbol{v}')^2}{c_0^2} + \mathcal{O}\left(\frac{v^4}{c_0^4} \right) \right) , \tag{25}$$

where $\boldsymbol{v}'$ is the velocity of the clock with respect to the laboratory. Again, α vanishes in SR. Equations (24) and (25) require three independent test in order to fix A, B, and α. As a prerequisite, the Robertson-Mansouri-Sexl test theory requires the independence of the speed of light from the velocity of the source.

In dynamical test theories of the Ni–Haugan–Kostelecky type the velocity of light again depends on the orientation and a velocity. The velocity of light can be calculated from the dispersion relation resulting from the modified Maxwell

[4] However, at least for the existing kinematical test theories one nevertheless needs some physical information "from the outside", namely the velocity with respect to some preferred frame. What we take as preferred frame depends on our knowledge about the universe. These test theories are not intrinsically complete.

equations (17)

$$\omega = \left(1 - \frac{1}{2}\tilde{k}_\alpha{}^\alpha \pm \sqrt{\frac{1}{2}\tilde{k}_{\alpha\beta}\tilde{k}_{\alpha\beta} - \frac{1}{4}(\tilde{k}_\alpha{}^\alpha)^2}\right)|\boldsymbol{k}| + \mathcal{O}(\lambda^2)\,, \tag{26}$$

where $\tilde{k}^{\alpha\beta} = \lambda^{\alpha\beta\gamma\delta}k_\alpha k_\beta/|\boldsymbol{k}|^2|$ [43]. The velocity of light then is given by

$$c(v,\vartheta) = c_0\left(1 + A\sin\vartheta + B\cos\vartheta + C\sin(2\vartheta) + D\cos(2\vartheta)\right)\,, \tag{27}$$

where the coefficients A, ..., D depend on the velocity with respect to an underlying coordinate system which in this case is chosen as the coordinate system of the solar system [43] (see also [44] for an application of this to the analysis of a recent experiment). The more complicated orientation dependence is due to the tensorial character of the LLI violation λ^{abcd}. For the time dilation one needs also to calculate the influence of the modified Maxwell equations on the atomic spectra. In this test theory, too, the independence of the speed of light from the velocity of the source is a prerequisite.

Test of Universality of *c*. The universality of the maximum velocity of particles has been tested for various pairs of particles. For a comparison of electrons with photons and of photons with different frequencies in the laboratory one achieves the 10^{-6} level for the relative velocity difference $|(v_1 - v_2)/v_1|$ [82–84]. Astrophysical observations of neutrinos and photons from supernovae gives a level of 10^{-8} [85–87]. High energy cosmic rays are discussed in [88–90]. Another aspect of the universality of c is the non-occurence of birefringence: in vacuum the velocity of light should be independent of the polarization. This has been confirmed by astrophysical observation with high accuracy [43] from which 10 of the 19 components λ^{abcd} could be estimated to be smaller than $2 \cdot 10^{-32}$ (the other 9 components will be constrained by Michelson–Morley experiments). More indirect nuclear spectroscopical (Hughes–Drever) tests give a 10^{-22} level for the maximum difference of photon and proton speed [91,92].

Test of Independence of *c* from the Velocity of the Source. One of the most distinctive and contra-intuitive statements of SR is that the velocity of light is independent from the velocity of the source. A possible violation may be expressed as $c' = c + \kappa v$ where κ is a parameter which has to be determined experimentally. In a Galilean framework $\kappa = 1$, in SR $\kappa = 0$. The most impressive experiment demonstrating this is from Alväger *et al.* [93] where the source of photons possesses a velocity of 99.975 % of the velocity of light. The emitted photons still propagate with the velocity of light within $\kappa \leq 10^{-6}$. Better estimates can be achieved from astrophysical observations of binary systems [94] leading to a $\kappa \leq 10^{-9}$.

Test of Isotropy of *c*. The isotropy of the velocity of light is subject to the famous Michelson–Morley experiments. The presently most precise test has been performed by Müller *et al.* [45] and gives $|\delta_\vartheta c/c| \leq 4 \cdot 10^{-15}$. In terms of the Robertson–Mansouri–Sexl parameters this means $|B| \leq 4 \cdot 10^{-9}$ and for remaining 9 parameters of the extended standard model $|\lambda^{abcd}| \leq 10^{-15}$. Other recent experiments are [95,44].

Since light, which properties are tested in these experiments, is a consequence of Maxwell's equations, any modification in the speed of light must be related to a modification of Maxwell's equation. Since the properties of the interferometer arms or the resonators are also strongly influenced by electrodynamics, their porperties may be modified, too. Indeed, it has been shown in two models [47,45] of modified Maxwell equations that an accompanying anomalous behaviour of the length of the interferometer arm or of the cavity may compensate or enhance the signal indicating an hypothetical anisotropy of the speed of light. Furthermore, also LLI-violations of the Dirac equation may contribute to an anomalous behaviour of the interferometer or cavity [96]

Test of Independence of c from the Velocity of the Laboratory. This part of the relativity principle has been tested first by Kennedy and Thorndike. The presently most precise test is due to Wolf *et al.* [95] and gives $|A| \leq 3.1 \cdot 10^{-7}$. Another recent experiment is [68][5].

Test of Time Dilation. Time dilation is the only non-null test of SR because one has to determine the Lorentz factor $1/\sqrt{1-v^2}$ experimentally. Deviations from this factor in terms of a parameter α in $\gamma(v) = (1 + \frac{1}{2}\alpha v^2)/\sqrt{1-v^2}$ have ben limited to $|\alpha| \leq 2 \cdot 10^{-7}$ recently [98].

6 New Experimental Devices and Developments

Here we describe a few experimental devices that have been developed in the last years and possess the capability to contribute a lot to improvements of experiments searching for new physics.

6.1 Atom Interferometry

Though atomic iterferometry has been implemented only a bit more than ten years ago, it already provides e.g. the best gyroscopes. High precision atomic interferometry is based on an effective laser cooling of atoms down to temperatures

[5] It should be noted that in terms of the variation of the velocity of light for varying velocities of the laboratory, $\delta_v c/c$, the old 1990 experiment by Hils and Hall [97] is better than the present tests. The difference is that Hils and Hall were able to measure for a few days only, while the measurements of Braxmaier *et al* [68] took approximately one year. Consequently, the change in the velocity which is essential in estimating the parameter A in (24), could be chosen as twice the velocity of the Earth around the Sun while Hils and Hall were restricted to twice the rotational velocity of the Earth's surface around its own axis.

in the μK domain corresponding to velocities of the order of cm/s. These low velocities are necessary for long interaction times thereby increasing the accuracy. The use of laser beams as beam splitters has the advantage of being not influenced by the gravitational or inertial field, as it is the case for the beam splitters in neutron interferometry. Within an Newtonian framework, acceleration and rotation gives as phase shift in an atom interferometer

$$\delta\phi = \boldsymbol{k} \cdot \boldsymbol{g}\, T^2 + \boldsymbol{k} \cdot (\boldsymbol{\Omega} \times \langle \boldsymbol{v} \rangle) T^2 \,, \tag{28}$$

where $\boldsymbol{k}$ is the wave vector of the laser field, $\boldsymbol{g}$ the (gravitational or inertial) acceleration, $\boldsymbol{\Omega}$ the angular velocity, $\langle \boldsymbol{v} \rangle$ the expectation value of the velocity of atoms entering the interferometer and T the interaction time. Therefore, the UFF is clearly represented by this phase shift [99]. If the inertial and gravitational mass are different, then this first term will be modified to $(m_{\mathrm{g}}/m_{\mathrm{i}})\boldsymbol{k} \cdot \boldsymbol{g}\, T^2$. Atomic interferometry has confirmed the UFF in the quantum domain to the order of 10^{-9}. It is astonishing that (28) is an exact quantum result though there appears no $\hbar$ in it.

Further improvements are expected by using Bose–Einstein condensates as a coherent source for atoms.

6.2 Atomic Clocks

There are various atomic clocks available: H-maser, Rb- and Cs-clocks, and ion clocks based on Hg^+, Yb^+, or Cd^+, see e.g. [100,101] for reviews. The accuracy of a clock is based essentially on the line-width of the atomic transition and on the time of interaction with an external oscillator which reads out the frequency. A narrow line-width is related to long-living atomic states that are provided by hyperfine states. These transitions possess frequencies in the microwave range (several 10 GHz). The interaction time should be 1 ms or longer.

Atomic Clocks. Clocks like the conventional Cs atomic clock consist of an atomic beam which, during its flight, is interrogated by some microwaves for a certain interaction time. Due to gravity, a relatively high beam velocity must be chosen, so the interaction time is limited to about 1 ms. Due to this limitation and further unwanted effects like stray fields, Doppler broadening, etc., the accuracy of such clocks is of the order 10^{-14}. For a detailed discussion, see [101]. The today's definition of the second is based on the Cs clock.

H-Maser. H-Masers are based on the coupling of the hyperfine transition of the ground state of the Hydrogen atom which has a lifetime of about 1 second, to the radiation of a resonator. The frequency is 1.420 405 751 Hz and the instability of this clock is of the order 10^{-15}. H-Masers are used worldwide for the definition of time and have been used in the first gravity space mission GP-A [64,65].

Ion Clocks. Ion clocks are also based on hyperfine transitions of ions which are stored in traps and which are therefore isolated from many disturbing influences. The instability of this kind of clocks approaches the 10^{-16}-level [102]. This means that within 1 billion years the clock may be wrong by 1 s. Ion clocks may be used in space missions like SPACETIME [70,69] or OPTIS [71].

Atomic Fountain Clock. Atomic fountain clocks use laser-cooled atoms. During a ballistic flight these atoms interact with separated fields in a Ramsey set-up. Due to the controlled flight of the atoms interaction times of up to 1 s on Earth can be obtained. This can be increased considerably in space where the atoms do not fall out of the apparatus. The corresponding project PHARAO on the ISS [103,69] is near completion.

6.3 Ultrastable Cavities

Cavities are made of stable materials and define a length standard. This lenght is read out by a laser frequency stabilized ("locked") to a resonance of the cavity. With $\nu = nc/L$ the information of the length L of the cavity is transformed to a frequency which can be measured with higher accuracy than lengths. Consequently, resonators are a realization of light clocks. Here, the velocity of light c plays an important role. It is clear that it is crucial to prohibit the cavity from any thermal expansion or distortions due to external forces like acceleration, rotation or gravity gradients. Furthermore, precise control of the laser frequency to the resonance frequency of the cavity is required. This can be controlled by means of the so-called Pound-Drever-Hall technique where one measures the modulation of the reflected or transmitted beam that.

For cryogenic optical resonators [104] the stability which can be achieved is $\delta L/L \leq 6 \cdot 10^{-16}$ [105] what, for a resonator of typical length of 5 cm, is about $1/100$ of the radius of the proton.

6.4 Frequency Comb

Since most of the tests of the principles of relativity depend on clock comparison, a high precision technique for comparing frequencies of various ranges is mandatory. For a comparison of microwave and optical frequencies, which differ by up to 6 orders of magnitude, the recently invented frequency comb is the appropriate technique, see [106] for an overview. This technique, being simpler, cheaper, and more accurate than previous methods, can be used e.g. in Kennedy–Thorndike tests and test of the UGR. Corresponding tests are under development at the University of Düsseldorf.

7 EEP and Modern Metrology

Metrology is the definition, preparation, transport, and comparison of physical units like the second, the meter and the kilogram. It can be viewed as the basis for all modern physics: without the possibility to make very precise measurements no progress in physics is conceivable. For example, in order to prove the predicted dynamics of a certain physical phenomenon, a precise time-keeping is required.

Since all physical units are represented themselves by some physical phenomena, a measurement always consists of the comparison of two physical phenomena of the same kind, like e.g. the measurement of the dynamics of the Earth compared with the dynamics of an atomic clock. The important point is therefore the *reproducibility* and *stability* of physical units. The reproducibility of units based on quantum effects is arbitrarily good while the reproducibility of a macroscopic meter stick or of a unit of mass is of the order 10^{-6} and therefore not applicable for really high precision measurements. In fact, the uncertainty in the definition of the mass unit (and in the homogeneity of the used masses) is the main obstruction for precise measurements of the gravitational constant. However, basing units on quantum effects also means that one relies on a certain structure of quantum mechanics. As we have seen, this structure is deeply related to the validity of the EEP.

7.1 Ideal Rods and Clocks

In order to explore the laws of physics and to perform tests of foundations of theories one has to measure or prepare certain quantities like time and length. The first step always consists of the definition of certain quantities like the second or the meter at a certain instant and at some position. The next step then is to transport this unit to other places. A complete theory always defines a way to transport these units. For example, within SR and GR one can design a transport of length and time standards with light rays and freely falling particles only. It has been shown in the axiomatic approach to GR (and thus also to SR) using light rays and freely falling particles (those obeying the UFF) only [11], that by means of Schild's ladder [11,107] or Perlick's construction [108] it is possible to uniquely transport a length or the eigentime along a path. For more general theories the uniqueness will be lost. In a Weylian model of gravity, for example, the transport of a length scale depends on the path and thus on the history of, e.g., the meter stick, see e.g. [108]. As a consequence, modern metrology with its task of unique definition, reproduction, and transport of physical units is deeply connected with SR and GR.

Another point is that though the above constructions need no other objects than those given by the theory, these transport prescriptions of time and length standards are not always practical procedures because they may not be realizable with the accuracy needed today. The length and time standards, which are realized today with the highest internal accuracy, are provided by atomic clocks and solids. In order to describe these standards, one needs more than just light rays and particles, namely the equations of motion for electromagnetic and quantum fields.

Another definition and transport of a certain length scale is provided by quantum equations for massive particles, e.g. the Dirac equation. The Dirac equation introduces the Compton–wavelength and its transport along quasiclas-

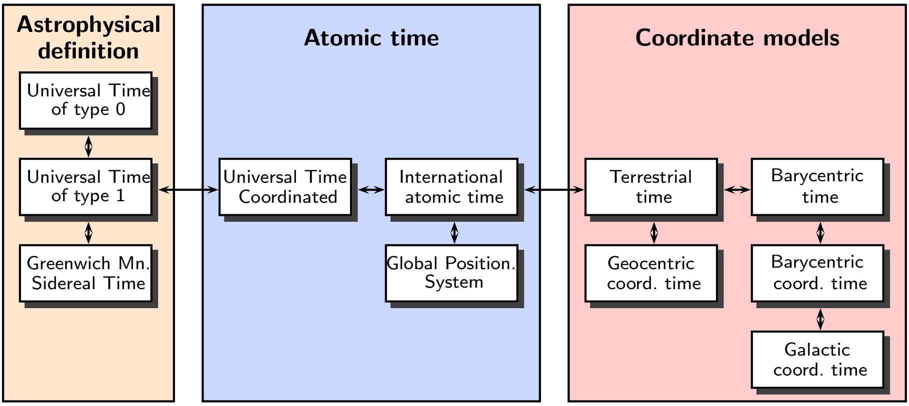


Fig. 3. Time scales and their relations. For all relations between the various time scales but the astrophysical ones, one needs SR and GR. See [109,110]

sical trajectories. Though the reproducibility of this length standard is very high, the Compton wavelength cannot be related to frequencies with high accuracy. Therefore, also this ideal standard is again not a practical one.

From these examples it is also clear that the unique availability of physical units is a matter of the physical *dynamics*. This is the case, because all physical standards are more or less defined by complicated physical objects that evolve in time according to the underlying physical laws.

7.2 The System of Units

The first worldwide accepted units of time, length and mass were provided by the revolution of the Earth around its axis resulting in the *Universal Time* scales (of various types according to the kind of averaging), by a meter stick realized as a metallic bar having the length defined as one ten-millionth of the distance between the pole and the equator, and a mass unit made, like the meter, of platinum-iridium. The definition of time suffers from the irregularities of the motion of the Earth (the length of the day, for example, increases by around 2 seconds every century compared with a time-scale defined by more stable Quartz oscillators). In 1956, the Ephemeris Time, based on the Earth's orbital motion around the Sun, was chosen as definition for time: The second was the 1/31 556 925.9747 part of the year 1900. The problems with the definitions of length and mass were that a direct comparison of these macroscopic prototypes is nor very accurate, and that the intrinsic stability of these prototypes are not known. Each material, for example, experiences some ageing. Indeed, there is an unexplained drift of the various mass prototypes during the last decades.

A first step in improving this system of units was to replace the Universal Time by atomic time. This was a natural development since atomic clocks were much more precise than the astrophysically defined time unit. Therefore, during the 17th General Conference of Weights and Measures in 1983 one defined the

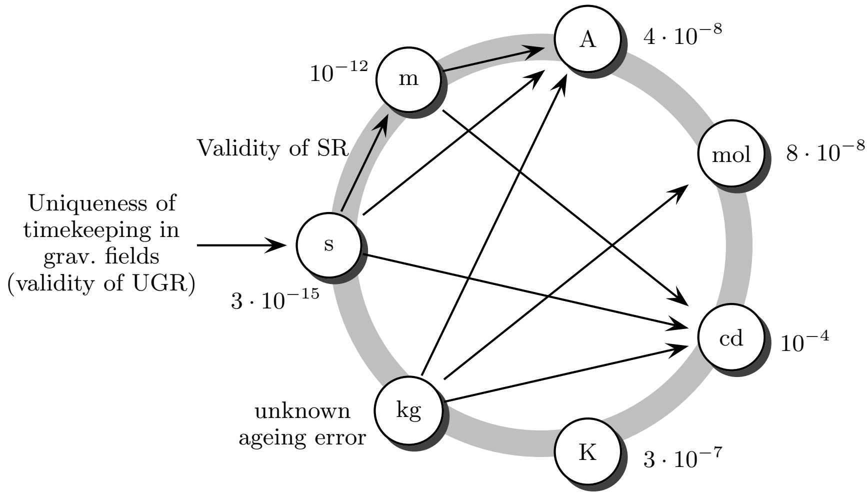


Fig. 4. The SI-units (s = second, m = meter, A = Ampere, mol = mole, cd = candelas, K = Kelvin, kg = kilogram) and their interdependences [111]. The numbers indicate the stability of the corresponding unit. The uniqueness of the transport of the definition of the second and of the meter depends on the validity of the EEP. A replacement of the mechanical definition of the Ampere or other quantities also requires conventional quantum theory and Maxwell theory and thus the validity of the EEP

second as the time that is needed by 9192631770 periods of the hyperfine transition of the ground level of ^{133}Cs. Therefore time was defined in terms of a highly reproducible quantum phenomenon. However, in order to combine all the atomic clocks around the Earth (which is advantageous since by enlarging the number of clocks, the precision of time-keeping will increase, and since astrophysical observations with distant telescopes, for example, need synchronization) Special and General Relativity is needed: We need Special Relativity for the synchronization procedure on a rotating reference frame (Sagnac effect), and we need General Relativity in order to account for different counting rates at different gravitational potentials (gravitational redshift). The result is the *terrestrial coordinate time* which is a model extracted from the reading of all the clocks on the surface of the Earth and which now represents the time in a non-rotating observer located at the center of the Earth. It is well known that time provided by the GPS also needs relativistic corrections.

The next step was to replace the unit of length by a much more reproducible phenomenon. In a first step this was done in 1960 by defining the meter as 1 650 763.73 wavelengths of the red $2\mathrm{p}_{10} - 5\mathrm{d}_5$ transition of Krypton. This length could be reproduced with an accuracy of $3 \cdot 10^{-10}$. The disadvantage was that the coherence length of that radiation was smaller than one meter, which made it difficult to be compared with the old standard. Later on, during the mentioned conference in 1983, the constancy of the speed of light was used in order to base the meter on time. Accordingly, the meter is now given as the distance light travels within the 299792458th part of a second. This can be realized very precisely using interferometry of laser beams. Therefore, with the help of Special Relativity, the length unit was replaced by a quantum phenomenon. Obviously, this definition breaks down if Special Relativity will be proven to be wrong, i.e., if the velocity of light depends on the orientation of the velocity of the laboratory or if the usual dispersion relation for light is modified.

The replacement of the definition of the kilogram by some quantum procedure is under way in various groups. One idea is to use a fixed number of atoms in a Si one-crystal. In principle, this is a well-defined, exactly reproducible procedure. However, counting the number of crystal lattices is not very practical so that one has to use optical techniques in order to measure the geometry of e.g. a sphere of a Si-crystal. That means that the kg will be based on the second. Another idea is to use the Watt balance which connects a mass unit and Planck's constant $\hbar$, and to replace the definition of mass by a definition of $\hbar$. Then $\hbar$ receives a *defined* value and the kilogram will be derived from it. Since the Watt balance relates mechanical power to electrical power with the help of the quantum Hall and the Josephson effect, again the Maxwell equations and the laws underlying quantum mechanics are involved.

A general task of modern metrology is to base all units on quantum mechanically defined and thus highly reproducible quantities. Beside the second and the meter, this has been done already for the electric resistance (measured in Ohm) and the electric potential difference (measured in Volt), based on the von-Klitzing and the Josephson-constant, respectively. Furthermore, the current can be based on the electronic charge and the second. Again, the validity of the standard Maxwell and quantum equations is a prerequisite for the consistent realization of this task.

7.3 Consequences of a Violation of the EEP

From this outline one can see immediately that the full system of units in its present (and proposed) form is compatible only if the present physical theories are correct, that is, if the EEP is valid. If, for example, SR is violated and the velocity of light is not isotropic, then the definition of length in general will not be unique in the following sense: For a given unit of time we may prepare two different length units in different directions. If we rotate these units of length, then they will in general not coincide if the velocity of light depends on the direction, that is, the unit of length prepared in the direction in which the velocity of light is larger will be smaller than the other one. One may think of an effect that internal forces of the material realization of the length standard may compensate for this effect. But one cannot expect this to happen for all materials in the same manner.

Another example is the uniqueness of the velocity of light: Let us assume that the velocity of light is different from the characteristic velocity appearing in the Dirac equation. Then the fine structure constant α as derived from spectroscopy by $\alpha = \sqrt{2\mathrm{Ry}\lambda_{\mathrm{C}}}$, where Ry is the Rydberg constant and λ_{C} the Compton

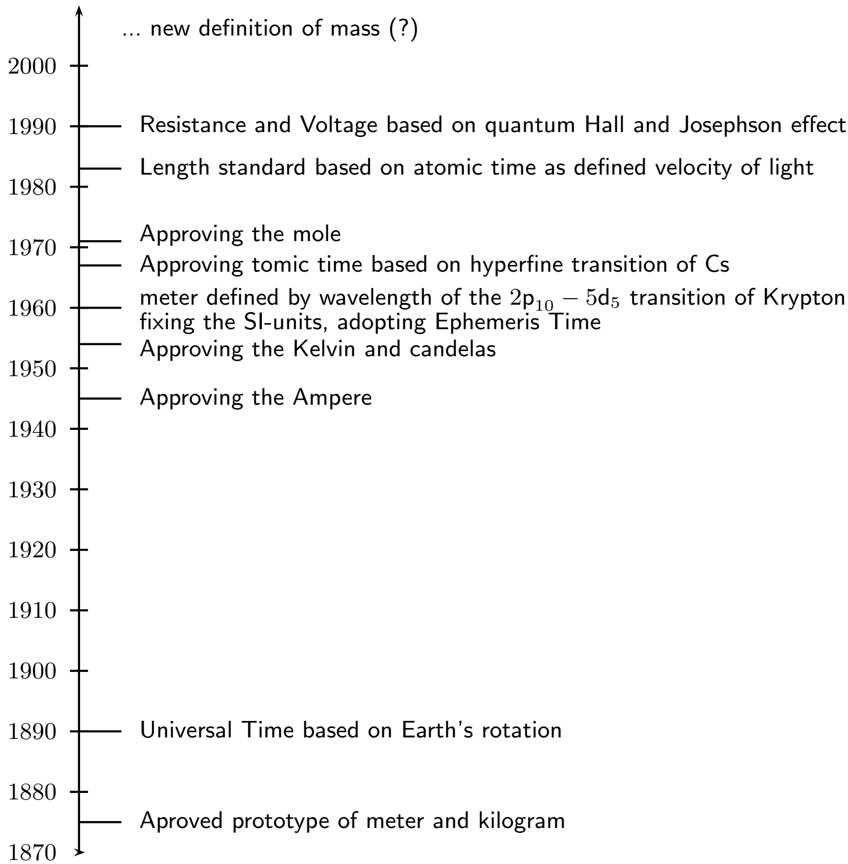


Fig. 5. Highlights of modern metrology

wavelength of the electron, will be different from the fine structure constant which can be derived from the quantum Hall effect by $\alpha = c/(2R_{\mathrm{K}})$ where R_{K} is the von Klitzing constant [112]. This again amounts to a non-uniqueness in the definition of length.

8 Conclusion

We showed that the implications of the Einstein Equivalence Principle are two-fold: First, the EEP strongly restricts the structure of the equations of motion for all types of matter fields, as for example the Dirac equation, and non-gravitational interactions like the Maxwell equations, and it fixes the structure of the gravitational field to be described by a metric field or a related quantity. The EEP is not sufficient to derive Einstein's field equation. Second, since the EEP determines the structure of standard physics, any deviation from standard physics should show up in violations of the EEP. Consequently, any search for violations of the EEP or any experimental improvement of the validity of the principles underlying EEP is very important for any theoretical scheme trying to go beyond the standard physics. Furthermore, the EEP also has very practical consequences in the sense that only for standard physics the today's scheme of metrology will give a consistent set of physical units needed to measure physical effects and compare theoretical predictions with experiments.

QG effects that fall outside the scope of the EEP and thus are not treated here, are

- Modifications of Newton's potential (see I. Antoniadis's lecture on page 337 and H. Abele's lecture on page 355).
- Time-dependent Newton's gravitational constant G, see e.g. [113].
- QG induced modifications of the dispersion relation leading to the GZK-cutoff presently very much discussed in astrophysics, see e.g. [114,115].
- QG induced noise in interferometers [116].
- QG induced decoherence in quantum matter, e.g. [117–119].
- QG induced fluctuation of the light cone [32].

Acknowledgements

I would like to thank Ch. Bordé, H. Dittus, H. Müller, and S. Schiller for stimulating discussions. Part of this article has been prepared at the Institute for Experimental Physics of the Heinrich-Heine-University Düsseldorf. I thank S. Schiller and his group for the hospitality and the inspiring atmosphere.

References

1. C. Kiefer: Quantum gravity – a general introduction. This volume
2. C.J. Isham: Conceptual and geometrical problems in quantum gravity. In Mitter H. and Gaustere H., editors, *Recent Aspects of Quantum Fields*, page 123. Springer Verlag, Berlin, 1991
3. C.J. Isham: A new approach to quantizing space-time: I. quantizing on a general category. gr-qc/0303060
4. C.M. Will: *Theory and Experiment in Gravitational Physics (Revised Edition).* Cambridge University Press, Cambridge, 1993
5. H.C. Ohanian: What is the principle of equivalence? *Am. J. Phys.* **45**, 903 (1977)
6. K. Nordtvedt: Space-time variation of physical constants and the equivalence principle. gr-qc/0212044
7. T. Damour: Gravity, equivalence principle and clocks. to appear in the Proceedings of the Workshop on the Scientific Applications of Clocks in Space (JPL, Pasadena, November 7-8, 1996)
8. T. Damour: Equivalence principle and clocks. In J. Trân Tanh Vân, J. Dumarchez, J. Reynaud, C. Salomon, S. Thorsett, and J.Y. Vinet, editors, *Gravitational Waves and Experimental Gravity*, page 357. World Publishers, Hanoi, 2000
9. G. Dvali and M. Zaldarragia: Changing α with time: Implications for fifth-force-type experiments and quintessence. *Phys. Rev. Lett.* **88**, 091303 (2002)
10. S. Sonego and V. Faraoni: Coupling to the curvature for a scalar field from the equivalence principle. *Class. Quantum Grav.* **10**, 1185 (1993)